A Regime-Based Evaluation of Southern and Northern Great Plains Warm-Season Precipitation Events in WRF

JINGYU WANG

Pacific Northwest National Laboratory, Richland, Washington

XIQUAN DONG

Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, Arizona

AARON KENNEDY AND BROOKE HAGENHOFF

Department of Atmospheric Sciences, University of North Dakota, Grand Forks, North Dakota

BAIKE XI

Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, Arizona

(Manuscript received 6 February 2019, in final form 9 April 2019)

ABSTRACT

A competitive neural network known as the self-organizing map (SOM) is used to objectively identify synoptic patterns in the North American Regional Reanalysis (NARR) for warm-season (April–September) precipitation events over the Southern and Northern Great Plains (SGP/NGP) from 2007 to 2014. Classifications for both regions demonstrate contrast in dominant synoptic patterns ranging from extratropical cyclones to subtropical ridges, all of which have preferred months of occurrence. Precipitation from deterministic Weather Research and Forecasting (WRF) Model simulations run by the National Severe Storms Laboratory (NSSL) are evaluated against National Centers for Environmental Prediction (NCEP) Stage IV observations. The SGP features larger observed precipitation amount, intensity, and coverage, as well as better model performance than the NGP. Both regions' simulated convective rain intensity and coverage have good agreement with observations, whereas the stratiform rain (SR) is more problematic with weaker intensity and larger coverage. Further evaluation based on SOM regimes shows that WRF bias varies with the type of meteorological forcing, which can be traced to differences in the diurnal cycle and properties of stratiform and convective rain. The higher performance scores are generally associated with the extratropical cyclone condition than the subtropical ridge. Of the six SOM classes over both regions, the largest precipitation oversimulation is found for SR dominated classes, whereas a nocturnal negative precipitation bias exists for classes featuring upscale growth of convection.

1. Introduction

In recent years, convection-allowing models (CAMs) have been extensively used in both operational and research environments to resolve the issue of convection in the coarse resolution of numerical weather prediction (NWP) models. With a horizontal grid spacing ≤ 4 km, convection is explicitly allowed and these high-resolution models can better represent the evolution and structure of mesoscale convective systems (MCSs; Weisman et al. 1997) than coarse-resolution models with parameterized convection (Kain et al. 2008). This characteristic makes CAMs appealing from numerous perspectives including operational forecasting, historical case studies, and climate downscaling. While the number of CAMs now being run is too long to list, efforts run in support of the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) Spring Forecast Experiment (SFE) include a long-term, deterministic WRF CAM from 2007 to present (hereafter NSSL-WRF), and the implementation of multiinstitutional ensembles such as the Community Leveraged Unified Ensemble (CLUE; Clark et al. 2018) in 2016.

Corresponding author: Xiquan Dong, xdong@email.arizona.edu

DOI: 10.1175/WAF-D-19-0025.1

Regardless of whether simulations are deterministic or ensemble in nature, the detailed structures seen in CAMs require careful considerations when applied to forecast verification. As theoretically demonstrated by Baldwin et al. (2001), a high-resolution forecast model that has a storm spatially displaced may perform worse than a coarse-resolution forecast for traditional verification methods (Jolliffe and Stephenson 2011; Wilks 2011). This result has also been observed in a number of model simulations. Using the University of Washington version of the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5), Colle and Mass (2000) and Mass et al. (2002) investigated precipitation in the Pacific Northwest. They found that traditional verification methods had large increases in accuracy as the grid spacing was decreased from 36 to 12 km, but this rate of improvement diminished as grid spacing was further reduced to 4 km. This result has also been seen for larger domains including the central United States using the NCEP Eta (Gallus 2002) and WRF (Done et al. 2004; Davis et al. 2006; Clark et al. 2007) Models. The main reason for the worse performance of the traditional evaluation methods in high-resolution models than their coarse-resolution counterparts is the ''double-penalty effect.'' As a result, newer verification methods are needed for the quantitative evaluation of the high-resolution models simulating convective features.

This issue is of special concern for regions such as the Great Plains (GP) that receives the majority of its annual precipitation during the warm season, with up to 60% of that total connected to MCSs (Ashley et al. 2003). From the perspective of the large-scale circulation, the baroclinic structure induces strong water vapor flux convergence over the central United States, resulting in a high frequency of convection (Wang and Chen 2009). Downscaling to the mesoscale, the low-level jet (LLJ) plays an important role by transporting large amounts of heat and moisture from the Gulf of Mexico northward (Weaver and Nigam 2008), whose interaction with large-scale synoptic patterns and subsequent impact on GP warm season precipitation has been categorized into two types: one associated with the upper-level synoptic trough, and the other tied to North Atlantic subtropical high (Weng 2000).

For the central United States including the GP, previous broad evaluations of precipitation have been performed for different CAMs (e.g., Kain et al. 2008, 2010a,b; Lean et al. 2008; Roberts and Lean 2008; Weisman et al. 2008; Clark et al. 2011) with different emphases on precipitation characteristics including timing, duration, evolution, distribution, and probability. Several studies have also investigated precipitation for the deterministic NSSL-WRF simulations discussed herein. Herman and Schumacher (2016) examined extreme precipitation events over the continental United States (CONUS) and found that most skillful prediction was associated with less extreme events (lower return period) while performance was highest from 1200 to 1800 UTC. Regarding the convection initiation process, a systematic eastward shift in NSSL-WRF forecasts was found by Coffer et al. (2013), and this issue was more prominent for cases with propagating, large-scale precipitating systems. In addition to the conventional evaluation practices, more recent CAM evaluation works utilize the object-based approach that first identifies and tracks the MCS objects at both the simulated and observed fields, and then the precipitation and cloud properties are compared between the two MCS objects over their entire life spans (e.g., Clark et al. 2014; Feng et al. 2018).

By themselves, traditional or newer verification methods do not determine the physical reasons or causes of model deficiencies. Further, there is no guarantee that methods will produce similar results over all times, regions, or environmental factors as evidenced by variability in model performance (Fritsch and Carbone 2004; Done et al. 2015; Surcel et al. 2016; Dey et al. 2016). As a result, verification studies can be less useful to individuals interested in the performance of models for forecasting. Instead, alternative methodologies must be developed to understand current CAM performance and identify future efforts to improve these models. In Coffer et al. (2013) for example, manual identification of 500-hPa synoptic patterns determined biases were larger for drylines associated with synoptic-scale wave activity versus those in quiescent conditions. Goines and Kennedy (2018) investigated spatial and temporal properties of precipitation from a subset of the NSSL-WRF forecasts (2010–12) and noted a positive bias of precipitation associated with convection driven by diurnal heating. Over the Southern Great Plains (SGP) of the United States, this was offset by a nocturnal, negative bias, whereas biases were predominantly positive over the Northern Great Plains (NGP).

The segregation of model performance by meteorological regimes has been commonly used in the climate modeling community where ample data allow for separation of model performance by prevailing conditions or synoptic patterns. In turn, this can provide insight into model behavior (e.g., forcing mechanisms responsible for evaluated events). From an NWP perspective, this can also provide practical information to forecasters regarding environments associated with model biases. The complexity of regime-based analyses varies, ranging from manual selection (Lamb 1959) to simple segregation

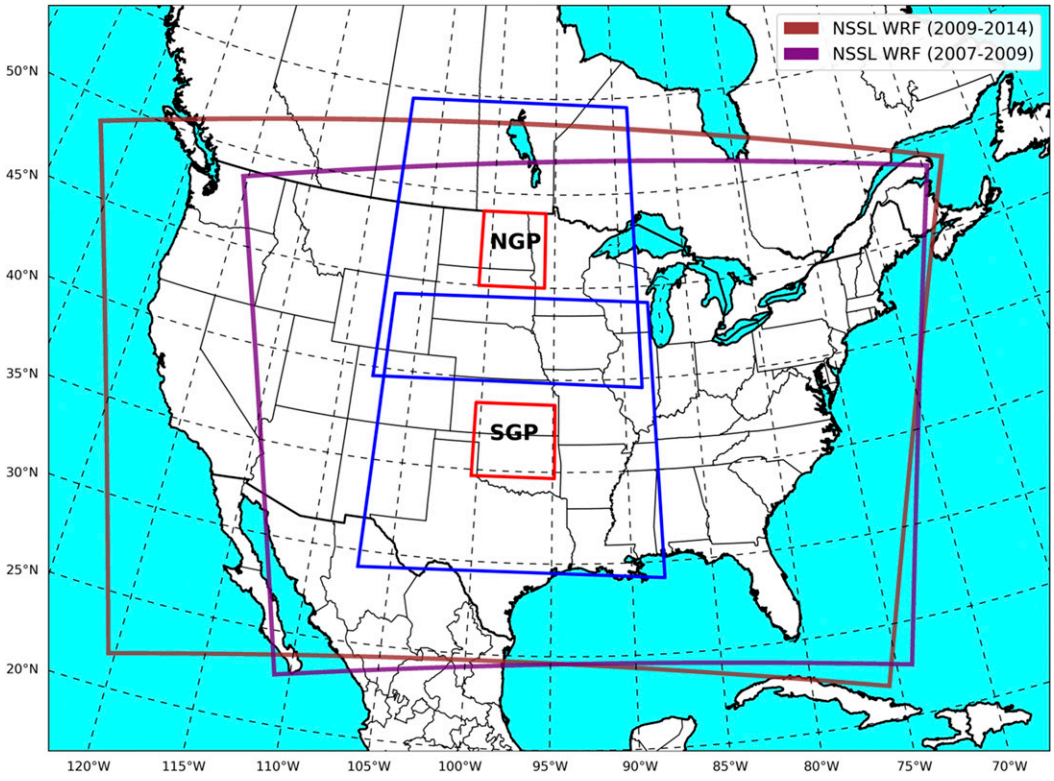

FIG. 1. NSSL-WRF domains and the NGP and SGP regions analyzed in this study. Blue boxes represent domains for the classification of meteorological patterns, and the smaller red boxes represent the areas precipitation was evaluated.

by modeled or reanalyzed vertical motion (e.g., Tselioudis and Jakob 2002) to clustering via neural networks (e.g., Marchand et al. 2006, 2009). Specifically over the GP, Chen and Kpaeyeh (1993) utilized the synoptic-scale environment to investigate the LLJ's climate impact. A ''fingerprint'' pattern recognition technique was adopted for major weather events identification (Root et al. 2007). Van Weverberg et al. (2018) used different cloud regimes to analyze the surface radiation biases in NWP at the Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) SGP site.

While a number of neural network clustering techniques are available, self-organizing maps (Kohonen et al. 1996) are now commonly used to type atmospheric and oceanic patterns as reviewed in Sheridan and Lee (2011) and Liu et al. (2006). Since these studies, an additional 50 articles in AMS journals mentioned SOMs in their titles. At the most basic level, the SOM is a *K*-means clustering algorithm with a neighborhood function included. This neighborhood function has two goals. First, it allows clusters to be related to each other in a two-dimensional sense, providing a visually intuitive way to analyze results. Second, it impacts the training process by weighting multiple clusters during the classification process. The result is a topological map with a range in clusters that span the original data space, sacrificing some exactness for a continuum of patterns. With patterns of more similarity closer together, less emphasis is placed upon the user knowing the ''correct'' number of classifications in advance. Rather, the user can investigate the topological map and make decisions on what nodes are meteorologically relevant.

The purpose of this study is to expand upon the earlier analysis of warm-season precipitation within the deterministic NSSL-WRF simulations by Goines and Kennedy (2018). Rather than focusing on specific events, the long-term nature of these simulations allows for an investigation of CAM performance by atmospheric patterns, allowing for an evaluation strategy that can be useful as a form of pattern recognition. This was achieved by using SOMs to identify meteorological states from April to September 2007–2014 over the NGP and SGP (Fig. 1). These regions were selected based upon the gradient in precipitation bias noted in Goines and Kennedy (2018). With latitudinal variations in mind, this study will quantitatively evaluate the NSSL-WRF simulated precipitation for these regions by classified meteorological states, to provide robust statistical information

with respect to large-scale synoptic patterns that can be of use to the broader modeling and forecast communities.

This paper is formatted as follows. In section 2, the NCEP Stage IV observations, NSSL-WRF simulations, and the SOM technique is introduced. This is followed by additional methodology to describe how NCEP Stage IV hourly precipitation data are separated into convective rain (CR) and stratiform rain (SR) components following a rainfall-rate criterion (RRC) method. In section 3, the results of the SOMs for the SGP/NGP are discussed, and NSSL-WRF simulations and Stage IV observations are analyzed under different classified atmospheric patterns. The analysis mainly focuses on the diurnal cycle, averaged precipitation intensity and coverage, and separation of CR versus SR over both regions. Four performance indices are developed for the quantitative evaluation of NSSL-WRF precipitation simulation under each SOM class. Finally, conclusions and suggestions for model improvement are discussed in section 4.

2. Data and methodology

a. NSSL-WRF simulation

NSSL has run a daily (0000 UTC), 4-km, deterministic, Advanced Research version of WRF (WRF-ARW) simulation in support of the SFE from 2007 to present. Integrated over 36 h, the simulations have 35 vertical levels and a time step of 24 s. Run as a singular domain, initial and lateral boundary conditions were provided by the North American Mesoscale Forecast System (NAM) interpolated to a 40-km grid. While the domain (encompassing most of the CONUS; Fig. 1) and core (currently v3.4.1) have been updated over time, a unique aspect of this set of simulations is a set of physical parameterizations that have had minimal changes. These include the WRF single-moment 6-class (WSM6) microphysics parameterization scheme (Hong and Lim 2006), Dudhia (Dudhia 1989) and RRTM (Mlawer et al. 1997) radiation parameterizations for shortwave and longwave radiation, and Mellor–Yamada–Janjić (MYJ) planetary boundary layer scheme (Mellor and Yamada 1982).

b. Precipitation observation

Given the high-resolution nature of NSSL-WRF simulations, precipitation was compared to the National Centers for Environmental Prediction (NCEP) Stage IV multisensory analysis (Baldwin and Mitchell 1997; Lin 2011; Lin and Mitchell 2005). Although lower in resolution (hourly, 4-km grid spacing) compared to the NSSL National Mosaic and Multi-Sensor Quantitative (NMQ) Precipitation Estimation (QPE, Q2) system (Zhang et al. 2011), manual QC of this product improves many aspects of precipitation estimates such as contamination by ground clutter (Lin and Mitchell 2005). Despite this QC effort, the NCEP Stage IV precipitation product is not flawless as pointed out by numerous studies (Schwartz and Benjamin 2000; Stevenson and Schumacher 2014; Nelson et al. 2016; Hitchens et al. 2013; Herman and Schumacher 2016). Goines and Kennedy (2018) also identified other issues including evidence of bright banding and beam blockage when precipitation was accumulated over sufficient time (e.g., multiple seasons). These latter issues were not noticeable over the domains used in this study.

c. Selection of regions of interest

To investigate the gradient in precipitation bias seen in NSSL-WRF (Goines and Kennedy 2018), regions investigated included points centered over the ARM SGP site in Lamont, Oklahoma (36.6°, −97.5°), and over eastern North Dakota (47.0°, −98.3°; Fig. 1). Meteorological patterns were typed for 15° × 19° (latitude × longitude) regions surrounding these grid points (blue boxes in Fig. 1). This domain allowed for pertinent synoptic to meso-α-scale features to be investigated, but minimized the impacts of higher terrain (Rocky Mountains) on reanalysis fields. Precipitation was analyzed for smaller regions (4° × 5°, red boxes in Fig. 1) centered within the larger domains. This was done to help avoid issues related to displaced precipitation in the CAMs and to ensure precipitation events were centered within the larger-scale patterns. Other analysis regions (e.g., 1° × 1° and 2° × 2.5°) were tested, but correlations between daily accumulated NSSL-WRF and NCEP Stage IV precipitation totals were highest for the larger 4°×5° regions (0.86 vs 0.66 and 0.78, respectively).

Since the primary goal of this work was to identify regime-dependent characteristics of precipitation in observations and model simulations, the analysis focused on precipitating days where domain-averaged, accumulated precipitation fell within the upper 90% of the cumulative distribution function (CDF) of this property. This step was taken to strike a balance between sample size (adequate precipitating days) and significance of impact (removal of drizzle, isolated storms, etc.). SOMs were also created for all patterns (precipitating or not) and higher CDF thresholds. These tests demonstrated that (i) NSSL-WRF largely produced precipitation during the correct atmospheric states and (ii) similar albeit stronger patterns were identified for higher thresholds.

CDFs were calculated from ~1370 days that observations and NSSL-WRF simulations were available during the warm-season (April–September) from 2007 to 2014. Precipitation threshold values of 3 and 2 mm resulted in 387 and 421 case days for the SGP and NGP, respectively. Days were defined as periods from 1200 to 1200 UTC to align the analysis with the diurnal cycle of convection. This also allowed for the evaluation of 0000 UTC NSSL-WRF simulations for forecast hours 12–36, safely removing early hours that may have been prone to spinup issues.

d. The self-organizing map technique

SOMs can classify datasets with any arbitrary amounts of dimensions into a 2D matrix named as the feature map. Through the iteration process, patterns of more similarity are clustered together, forming a continuous spectrum with the most dissimilar patterns found at the beginning and the end of the feature map. SOMs mainly capture nonlinearities in the input data and provide a visually intuitive way to interpret results (Kohonen 1989). In this study, SOMs were created to objectively classify atmospheric patterns for the two regions. Because of the limited domain of the NSSL-WRF, the North American Regional Reanalysis (NARR; Mesinger et al. 2006) was used as input in a fashion similar to Kennedy et al. (2016). Given that precipitation days were defined from 1200 to 1200 UTC, atmospheric patterns were classified from the midpoint time at 0000 UTC, which also aligned with the time period when convection frequently initiates over these areas. To produce a classification that could be quickly interpreted from model output and be useful to forecasters, input variables include selected variables within the lower and midlevels of the atmosphere. These variables are mean sea level pressure (MSLP), relative humidity (RH), and wind (u and v components) at 900 and 500 hPa, and 500-hPa geopotential height anomalies. The choice of 900 hPa avoided problems with established surface biases in NARR (King and Kennedy 2019) while the use of height anomalies avoided issues with classifications being biased by seasonal variability in geopotential heights (Kennedy et al. 2016). Variables were averaged from the 32-km grid spacing within NARR to a 1°×1° grid to focus on larger-scale features and to decrease the computational power needed to create the SOMs. With 8 total variables and a 15° × 19° region surrounding each site, input vectors to train the SOM were 2280 (15 × 19 × 8) elements long. All variables were normalized to a common scale to contribute equally to the SOMs. To increase the robustness of statistics and reduce the burden of analysis, the original 28-class SOMs were averaged into final 3 × 2 (6-class) SOMs in this study. A brief review of SOM technique and detailed methodology are presented in the appendix.

e. Separation of convective versus stratiform rain

Through the combination of the Next Generation Radar (NEXRAD) network and the Geostationary Operational Environmental Satellite system (GOES) observations, Feng et al. (2011) developed a hybrid cloud classification algorithm that objectively separates the convective systems into the components of convective core (CC), stratiform rain (SR), and anvil clouds (AC). The SR regions have the largest coverage of warm-season rainfall over the midlatitudes, while the CC regions (corresponding to the precipitation type of CR) account for the most intense precipitation (Cui et al. 2019). Feng et al. (2012) also found that the CR rain rate is almost an order of magnitude higher than SR, causing a surge in accumulated precipitation within a short time period and possibly resulting in flooding events. Differences in statistical characteristics of CR and SR have been investigated through a variety of datasets, including space-borne satellite observations [e.g., Tropical Rainfall Measuring Mission (Yang and Smith 2006); GOES (Behrangi et al. 2009)], ground-based radar observations [e.g., National Mosaic and Multi-Sensor Quantitative Precipitation Estimation (Stenz et al. 2014, 2016; Feng et al. 2011, 2012)], direct surface rain gauge measurements (Giangrande et al. 2014; Wu et al. 2013; Tao et al. 2013), and aircraft in situ measurements (Beard et al. 1986; Wang et al. 2015, 2016, 2018).

Since the differences between CR and SR are so obvious, to understand whether simulated precipitation and associated biases are SR or CR in nature, this study separates these categories within NCEP Stage IV by using a rainfall-rate criterion method, where CR grid points within hourly precipitation can be identified if greater than a certain threshold (e.g., 10 mm h^{-1} at 1-km spatial resolution, Tokay and Short 1996; Nzeukou et al. 2004; Giangrande et al. 2014). The 10 mm h^{-1} RRC threshold is suitable for hourly data with a spatial grid spacing of 1 km but cannot be directly applied to Stage IV data with 4-km grid spacing. Rosenfeld et al. (1990) suggested that the optimal convective rain rate cutoff should fall between 4 and 6 mm h^{-1} at this lower resolution. In this study, grid points in Stage IV data greater than 5 mm h^{-1} are identified as CR.

f. Construction of performance matrices

For quantitatively evaluating the precipitation simulation on a high-resolution Cartesian grid, the uniform verification method of comparing domain averages is insufficient. As a result, a series of correlations and correlation-based measures were developed to examine

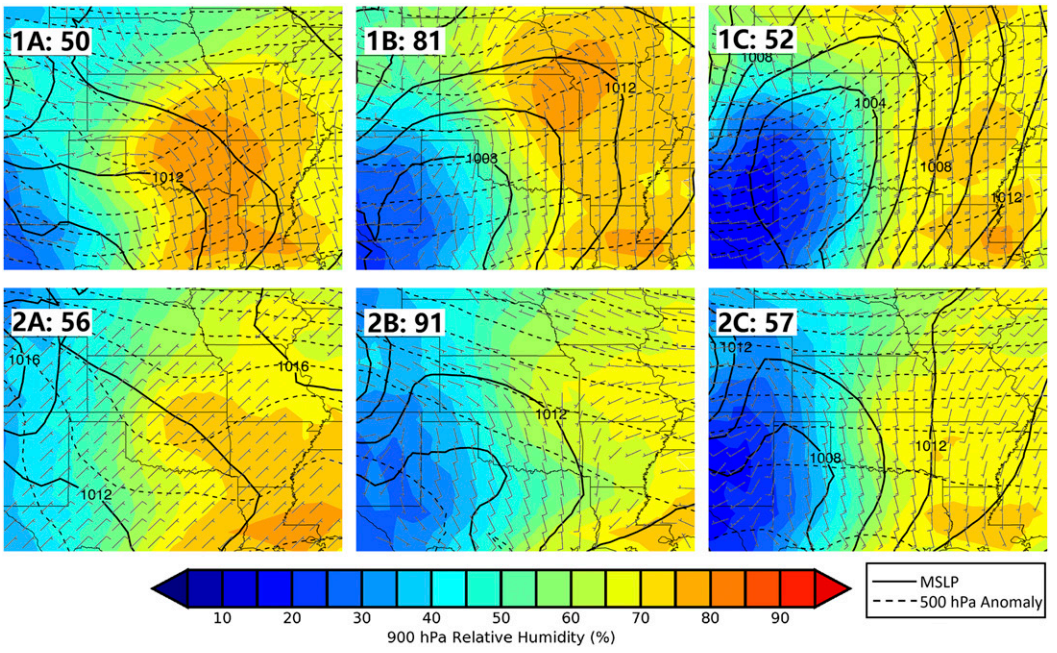

FIG. 2. Near-surface and midlevel analyses for the SGP SOM. MSLP is contoured with solid lines, 900-hPa RH with filled contours, and 500-hPa geopotential height anomalies with dashed lines. The 900-hPa wind barbs are also overlaid. Inset white boxes list the class number (1–6) and the number of cases per class.

the "goodness of fit" for hydrologic and atmospheric models (Legates and McCabe 1999), including correlation coefficient (CC), normalized standard deviation (NSTD), agreement index (AI), and ratio between simulation and observation:

$$R = \frac{\sum_{i=1}^{N}(S_i - \overline{S})(W_i - \overline{W})}{\sqrt{\sum_{i=1}^{N}(S_i - \overline{S})^2 \sum_{i=1}^{N}(W_i - \overline{W})^2}}, \quad (1)$$

$$\mathrm{NSTD} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(W_i - \overline{W})^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i - \overline{S})^2}}, \quad (2)$$

$$\mathrm{AI} = 1 - \frac{\sum_{i=1}^{N}(S_i - W_i)^2}{\sum_{i=1}^{N}(|S_i - \overline{S}| + |W_i - \overline{S}|)^2}, \quad \text{and} \quad (3)$$

$$\mathrm{ratio} = \frac{\sum_{i=1}^{N} W_i}{\sum_{i=1}^{N} S_i}, \quad (4)$$

where W_i and S_i represent the simulated and observed rainfall over each grid point, and a perfect simulation would be one in the above variables.

3. Results

a. The Southern Great Plains

1) Analysis of SOM patterns

The SGP is characterized by a variety of patterns responsible for precipitation events (Fig. 2). In the lee side of the Rocky Mountains, the majority of classes have a west–east gradient in humidity, representing drylines that are common in this region (Fig. 2). Classes 1A–1C (along the top row of the SOM) have southwesterly 500-hPa flow over the domain that transitions from a shortwave trough associated with a surface warm front (classes 1A–1B), to a stronger midlatitude cyclone underneath a larger-scale trough with a stronger 500-hPa jet streak (class 1C). In this latter class, the analysis domain fell within the warm sector ahead of the low pressure center. In many ways, these patterns are reminiscent of known patterns associated with severe weather and flash flood events for this region (Maddox et al. 1979; Nielsen et al. 2015), although it should be pointed out that the cases within this study simply represent the upper 90% of precipitating days. Along the bottom row (classes 2A–2C), 500-hPa flow

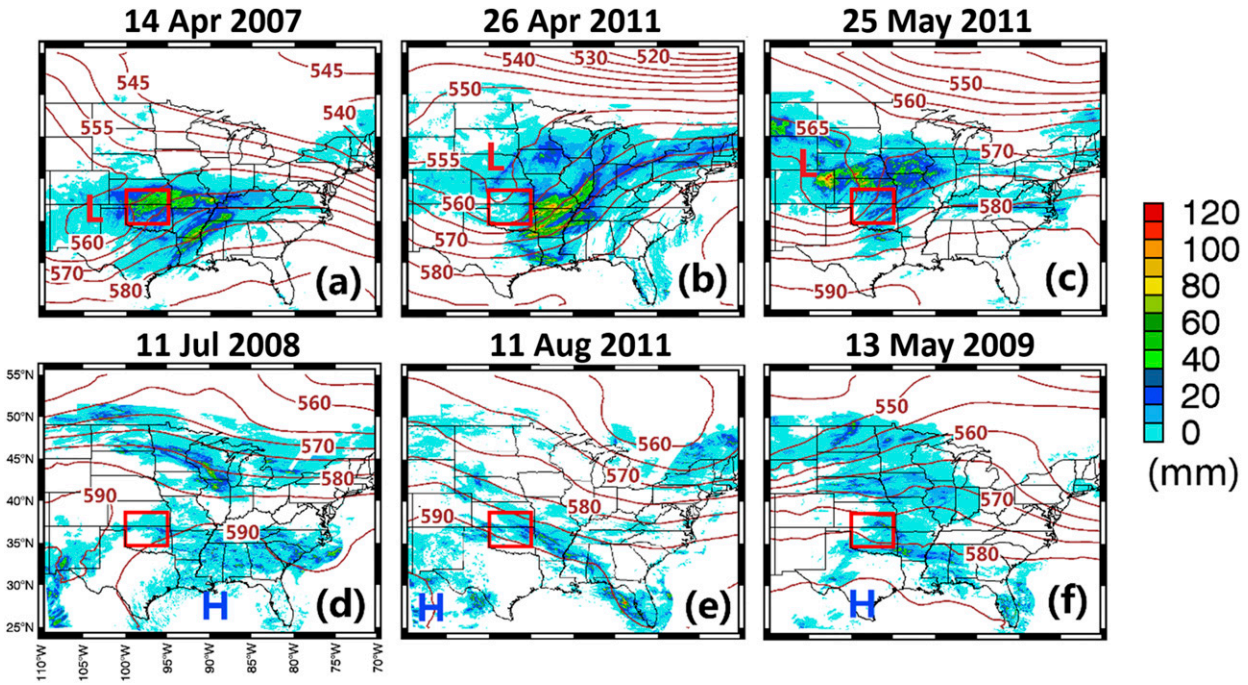

FIG. 3. Examples of precipitation cases for each class within the SGP SOM: (a) class 1A: 14 April 2007, (b) class 1B: 26 April 2011, (c) class 1C: 25 May 2011, (d) class 2A: 11 July 2008, (e) class 2B: 11 August 2011, and (f) class 2C: 13 May 2009.

transitions to weaker states with winds either northwesterly (classes 2A–2B) or zonal (class 2C). Contrasting classes 2A and 2B, the former category has a well-defined shortwave trough at 500 hPa suggesting stronger upper-level forcing. At the surface, conditions within classes 2A–2C are more quiescent with low pressure to the west, but with weaker pressure gradients and a dryline that strengthens from left to right (classes 2A–2C) across the SOM.

Examples of precipitation cases for each SGP class are shown in Fig. 3, where classes 1A–1C have a common feature of southwesterly wind with an upper-level trough or low in the vicinity of the SGP analysis domain. Evidence of the surface extratropical cyclone can be seen with the most intense precipitation occurring east of this feature, forming a widespread precipitation band over or near the study domain. As noted in previous studies (e.g., Wash et al. 1990; Schumacher 2017), these classes are most common in April to June and September, which is in line with the climatology of extratropical cyclones and statistics found in this study (Fig. 4). In the broadest sense, these patterns are tied to the location of the polar jet over this region. The polar jet stream plays a significant role for the cyclogenesis over the midlatitudes, especially from the east of the Rocky Mountains to the west of the analysis domain. As the jet stream intensifies in spring and early fall, a jet streak forms with upper-level divergence, which efficiently pumps air out of the vertical air column. In response to the divergence aloft, a low pressure system is generated at the surface.

Classes 2A–2B are characterized by northwesterly flow at 500 hPa over the west of the SGP domain and changed to near-zonal flow over the east of the SGP domain that is also occurred in class 2C (Fig. 3f). In addition to the difference in prevailing wind direction between the classes, differences in the morphology of daily accumulated precipitation are also apparent. The examples shown for classes 2A–2B have less intense precipitation around the periphery of a subtropical ridge/high center, which is commonly known as a ''ring of fire'' pattern (Galarneau and Bosart 2006). In the early efforts that synthesize heavy precipitation events with synoptic conditions (e.g., Maddox et al. 1978; Mitchell et al. 1995), the pattern of ''meso-high'' has been associated with severe weather. Dominated by the subtropical high pressure system, the air is most stable toward the center of the high pressure, where a subsidence inversion layer (capping inversion) is formed as a result of widespread descending air. The near-surface layer is heated and compressed by the high pressure but also trapped by the subsidence inversion, so the formation of thunderstorm is suppressed even with cold air aloft. However, this inversion becomes weaker toward the edge of the high pressure, which allows convection to occur with sufficient moisture supply, thus a ring of precipitation can form at

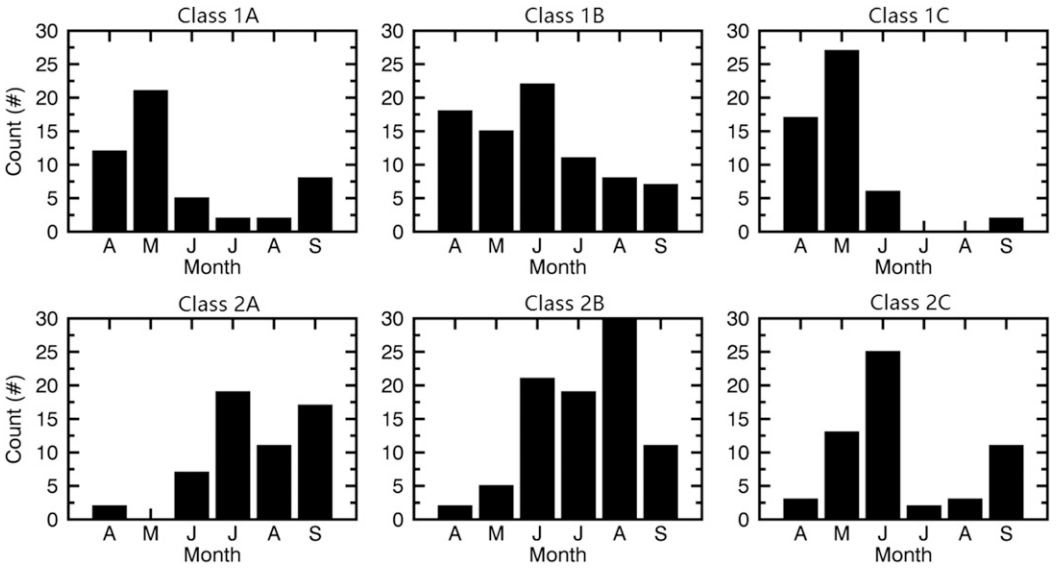

FIG. 4. Monthly number of cases per class for the SGP SOM displayed in Fig. 2.

the periphery of the high pressure. Similar to this bottom tier of classes, class 2C also has less intense precipitation, and there is some evidence of a weak shortwave trough at 500 hPa. Note that the original 28-class SOMs were averaged to the 3 × 2 (6-class) SOMs, the shortwave features are overwhelmed by averaging. As a result, the SOM only displays zonal flow for class 2C.

The preferred periods of occurrence of each SOM class are shown in Fig. 4. SGP precipitation cases are most common during the months of May–June with 81 and 86 events, respectively, from 2007 to 2014. Other months within the warm season had near-uniform counts with 53–57 events recorded. As noted earlier, classes with stronger surface and midlevel fields (e.g., classes 1A, 1C, 2C) have an increased likelihood of occurrence during the months of April-June and September in agreement with other studies that have classified atmospheric patterns over this region (Marchand et al. 2006, 2009). This also aligns with the climatology of severe weather over this region. Class 1B, which features the highest 900-hPa RH in association with a warm front also has an increased number of classes during these months although the patterns also occur from July to August. Classes 2A and 2B, which have northwest flow aloft over weaker surface patterns, are predominately summer patterns that are most frequent during June–September. Although surface patterns are less distinct in these classes, northwest flow events are known to produce severe weather episodes in this region during the summer (Johns 1984). Broadly speaking, one dimension of the SOM feature map (top vs bottom) captures the seasonal variability from spring to summer.

2) ANALYSIS OF THE DIURNAL CYCLE

The diurnal cycles of daily average precipitation rate over the SGP study domain are compared between Stage IV observations and NSSL-WRF simulations for the SOMs (Fig. 5). Whereas the *y* axis of the SOM demonstrates seasonal variability, the *x* axis (e.g., classes 1A/2A vs 1C/2C) shows an increase in the amplitude of the diurnal cycle. A two-peak distribution is found for classes 1A and 2A with a primary peak at sunrise (0600 LT) and the secondary peak at sunset (1800 LT). In contrast, the remaining classes have their peak rain rate around midnight, a typical diurnal pattern documented over the Great Plains (e.g., Kincer 1916; Wallace 1975; Colman 1990a,b). This region receives the majority of precipitation and convective activity at night during the warm season, which is not observed in other regions globally (e.g., Dai 2001; Nesbitt and Zipser 2003; Weisman et al. 2008). The lack of diurnal variation in classes 1A and 2A implies that the predominant forcing mechanism for these classes could be widespread SR precipitation. This argument can be supported by the synoptic patterns shown in Fig. 2, where these classes are associated with higher relative humidity, residing underneath a 500-hPa trough.

To visualize the contrast in diurnal precipitation variation, the daytime and nighttime averaged 12-h accumulated precipitation amounts from Stage IV observations and NSSL-WRF simulations, as well as their biases are shown in Fig. 6. The corresponding domain-averaged precipitation amounts are listed in Table 1. By comparing Figs. 6a and 6d, no significant difference is found

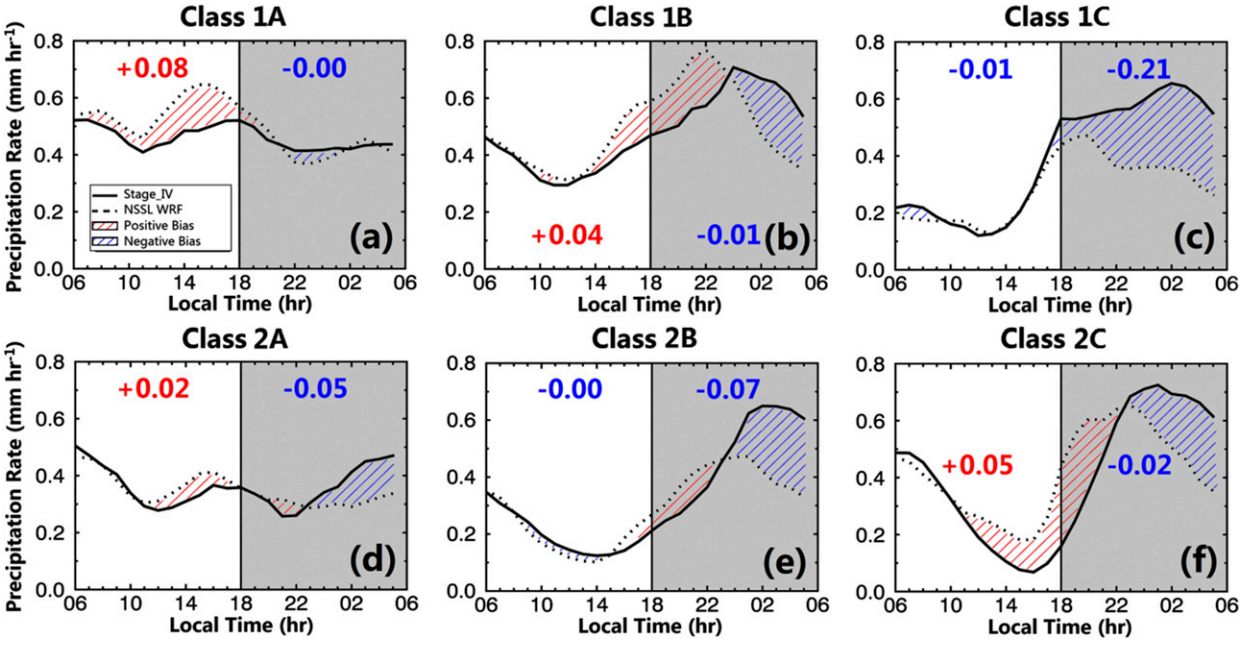

FIG. 5. Diurnal cycles of precipitation rate from Stage IV observations (solid lines) and NSSL-WRF simulations (dashed lines) for (a) class 1A, (b) class 1B, (c) class 1C, (d) class 2A, (e) class 2B, and (f) class 2C of the SGP SOM. The local night is shaded gray from 1800 to 0600 LT.

between the day and night observed precipitation for class 1A (5.74 vs 5.30 mm) and class 2A (4.38 vs 4.32 mm), whereas classes with more convective activity have lower daytime but higher nocturnal precipitation amounts (e.g., class 1B: 4.43 vs 7.09 mm; class 2C: 3.08 vs 6.61 mm). Moreover, the largest day/night contrast is found in class 1C (2.46 vs 6.96 mm) and class 2B (2.38 vs 5.55 mm), which indicates notable upscale growth at night (Figs. 5c and 5e). To reveal the significance of day–night contrast, the Bayesian estimation supersedes the *t* test (BEST; Kruschke 2013) method is performed among the observations of each class. The quantity examined is the differences between the observed nighttime and daytime mean precipitation. For robust statistical representativeness, the bootstrapping resampling is applied to enlarge the sample size. The *P*-value matrix (Table 2) shows that classes 1A and 2A are significantly different from the rest of classes ($P < 0.05$). This demonstrates that the well-documented GP nocturnal maximum precipitation diurnal pattern is not always true especially for events dominated by SR (classes 1A and 2A, taking up to 28% of all precipitating events). By comparing the precipitation patterns, the largest displacement is found in class 2C, where the heavy precipitation band is greatly shifted to the south in simulation. This shift is not seen for the rest of classes, which may reveal the uniqueness of this SOM pattern.

Comparing the NSSL-WRF simulations with Stage IV observations, regime- and time-dependent biases are seen (Fig. 5, Table 1). A negative nighttime bias is found for all classes, while a positive daytime/early evening positive bias is found for all classes except 1C and 2B. For class 1A, the nocturnal bias is negligible, and thus the model oversimulates precipitation over the entire day. In class 1B, the positive bias carries over into the evening, and this offsets a negative bias that begins at approximately midnight local time. This characteristic is common in the majority of the classes that demonstrate a large negative bias during the night, which was also noted in the study of Goines and Kennedy (2018). Even with explicit convection, proper simulation of the nocturnal maximum in precipitation remains a challenge in this CAM.

Through the examination of the BEST test results of the nocturnal negative biases in NSSL-WRF, the minimum *P* values are found associated with classes 1C and 2B featuring upscale growth in observation, and the former can pass the *t* test with significance level of 0.05 (Table 2). The test results indicate the issue of undersimulation in nocturnal precipitation is statistically significant for those classes.

The timing of the bias provides insight into the forcing mechanism. Focusing on the similarities between classes 1C and 2B, the peaks at 0200 LT correspond to the largest nocturnal negative bias. The earlier precipitation peaks in the NSSL-WRF simulations indicate that the

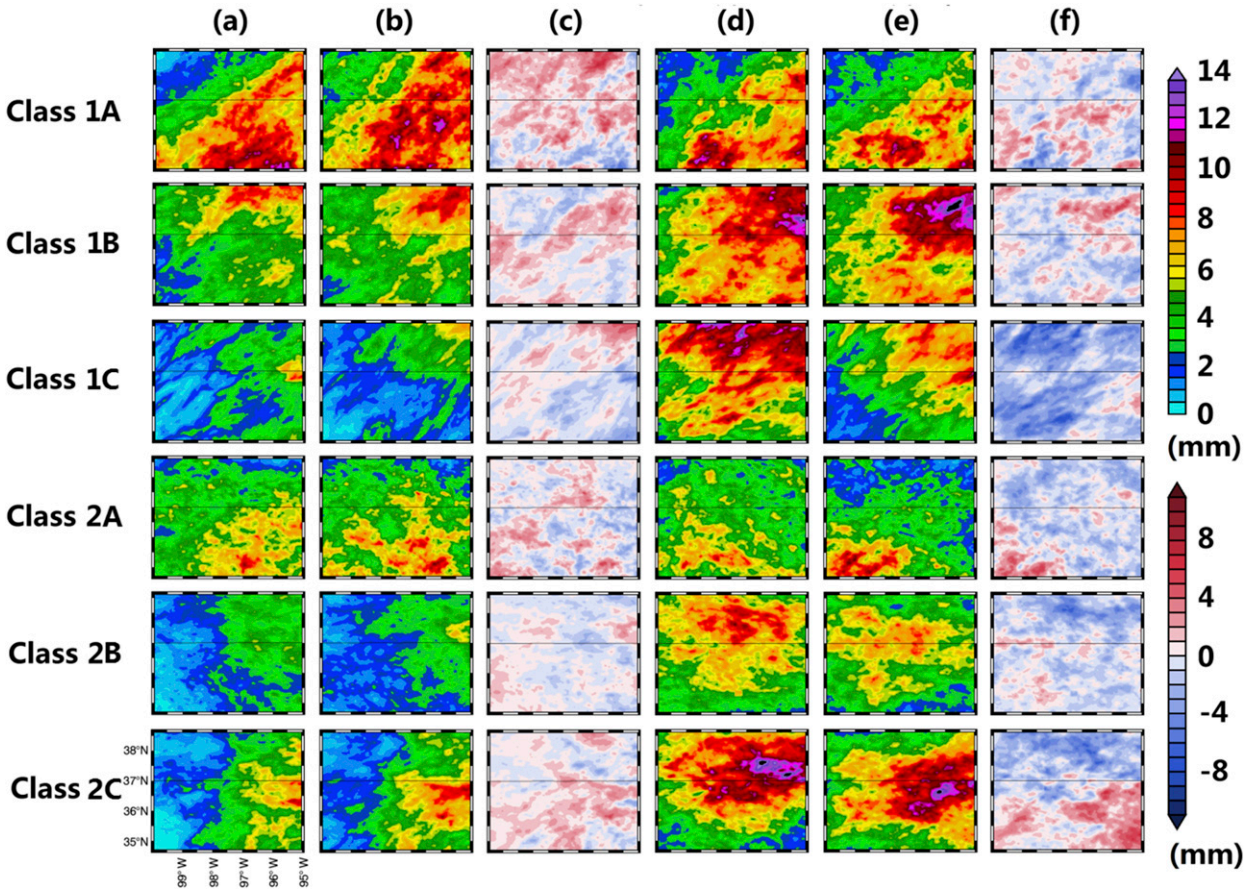

FIG. 6. Daily averaged, 12-h accumulated precipitation amount for (a),(d) Stage IV, (b),(e) NSSL-WRF, and (c),(f) bias for NSSL-WRF − Stage IV during the day in (a), (b), and (c) and night in (d), (e), and (f) for each class in the SGP SOM.

continuous upscale growth of convection during the night may be to blame, and this is supported by the Hovmöller diagrams shown in Goines and Kennedy (2018) that had propagating streaks that ended too early. Despite the negative biases of precipitation at night, classes 1C and 2B have excellent agreement during the day. That said, most of the classes (such as class 1B) have a well-defined phase shift with simulated

TABLE 1. Daytime and nighttime 12-h accumulated precipitation amount (domain average; mm) from Stage IV observations, NSSL-WRF simulations, their differences, and *t*-test scores for each class of the SGP and the NGP SOMs. The *P* values less than 0.05 are in bold.

			Class 1A	Class 1B	Class 1C	Class 2A	Class 2B	Class 2C
SGP	Day	Stage IV	5.74	4.43	2.46	4.38	2.38	3.08
		NSSL-WRF	6.72	4.92	2.34	4.60	2.35	3.65
		Difference	0.98	0.49	−0.12	0.22	−0.03	0.57
		P value	0.3904	0.5141	0.8463	0.7653	0.9393	0.4417
	Night	Stage IV	5.30	7.09	6.96	4.32	5.55	6.61
		NSSL-WRF	5.27	6.96	4.49	3.77	4.69	6.38
		Difference	−0.03	−0.13	−2.47	−0.55	−0.86	−0.23
		P value	0.9791	0.8944	**0.0129**	0.4599	0.2325	0.8559
NGP	Day	Stage IV	1.28	2.47	4.09	0.92	2.79	4.45
		NSSL-WRF	1.50	3.00	5.00	0.91	3.84	5.36
		Difference	0.22	0.53	0.91	−0.01	1.04	0.92
		P value	0.4916	0.2641	0.2310	0.9729	**0.0286**	0.2268
	Night	Stage IV	5.10	5.50	5.31	3.84	2.26	1.51
		NSSL-WRF	4.71	5.85	6.72	3.41	2.57	2.55
		Difference	−0.39	0.34	1.41	−0.44	0.31	1.04
		P value	0.6417	0.6208	0.1555	0.3777	0.4527	**0.0346**

TABLE 2. The *P*-value matrix of each SOM class's daytime/nighttime precipitation differences using the BEST method for the SGP and NGP.

		Class 1A	Class 1B	Class 1C	Class 2A	Class 2B	Class 2C
SGP	Class 1A	1	0.0121	0.0002	0.7215	0.0009	0.0102
	Class 1B	0.0121	1	0.1466	0.0156	0.6095	0.5305
	Class 1C	0.0002	0.1466	1	0.0001	0.2332	0.5383
	Class 2A	0.7215	0.0156	0.0001	1	0.0001	0.0111
	Class 2B	0.0009	0.6095	0.2332	0.0001	1	0.7738
	Class 2C	0.0102	0.5305	0.5383	0.0111	0.7738	1
NGP	Class 1A	1	0.3526	0.008 454	0.2366	3.473×10^{-9}	1.878×10^{-14}
	Class 1B	0.3526	1	0.04624	0.8785	2.180×10^{-7}	1.532×10^{-12}
	Class 1C	0.008454	0.04624	1	0.0431	0.0220	4.293×10^{-6}
	Class 2A	0.2366	0.8785	0.0431	1	1.941×10^{-8}	3.414×10^{-16}
	Class 2B	3.473×10^{-9}	2.180×10^{-7}	0.0220	1.941×10^{-8}	1	7.519×10^{-5}
	Class 2C	1.878×10^{-14}	1.532×10^{-12}	4.293×10^{-6}	3.414×10^{-16}	7.519×10^{-5}	1

precipitation peaking two hours early. The classes 1B and 2C serve as the transitional patterns connecting presumed SR precipitation (classes 1A and 2A) and those with nocturnal upscale growth (classes 1C and 2B).

3) ANALYSIS OF PRECIPITATION TYPE, INTENSITY, AND COVERAGE

As mentioned in section 2, the RRC method was used to segregate CR ($\geq$5 mm h^{-1}) and SR ($<$5 mm h^{-1}) grid points, and these data are averaged within each SOM class for intensity and coverage analysis (Fig. 7, Table 3). In this study, the precipitation coverage is calculated using the number of precipitating grid points divided by the total number within the domain. For all precipitating grid points, classes 1A and 2A have the lowest intensity but largest coverage that is primarily dictated by SR. Moving across the SOM, trends from classes 1A to 1C and from classes 2A to 2C are seen in these properties. The CR intensity and coverage for classes 1C and 2C appear to be maxima while SR coverages decrease to minima, indicating that convective activity plays a more important role in these two classes that grow upscale

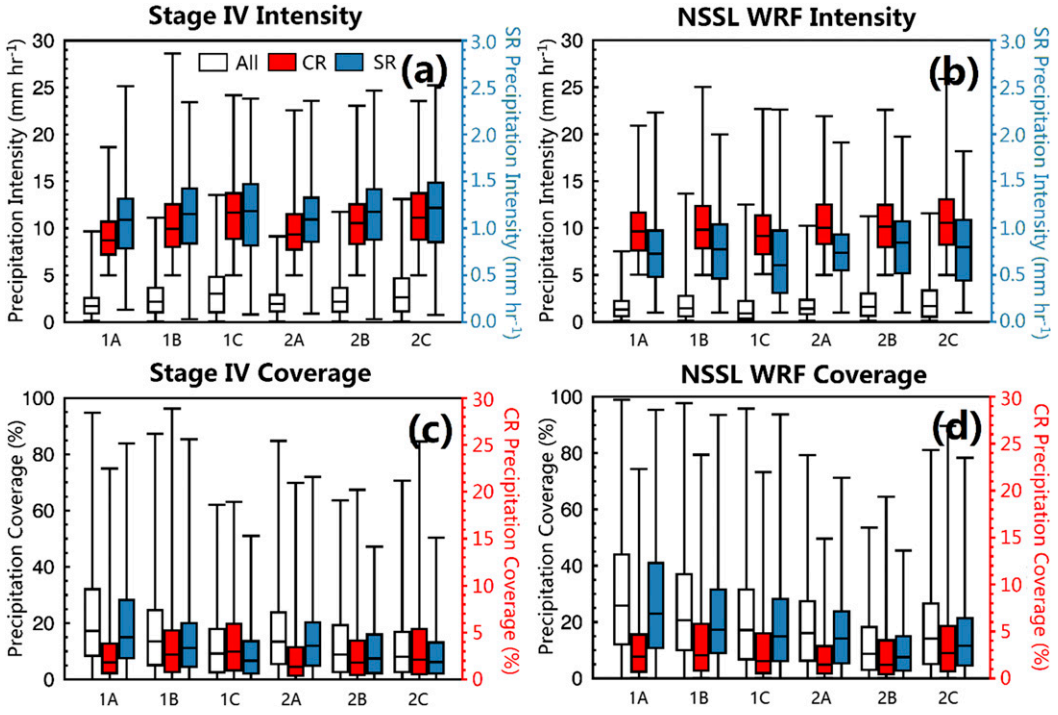

FIG. 7. Boxplots of (a),(b) precipitation intensity and (c),(d) coverage for (left) Stage IV and (right) NSSL-WRF for each class in the SGP SOM. Total precipitation is given by white boxes while CR is red, and SR is blue.

TABLE 3. Precipitation intensity and coverage from Stage IV and NSSL-WRF for each class of the SGP SOM. Precipitation is separated into total, CR, and SR components. The *P* values less than 0.05 are in bold.

	Precipitation intensity (mm h^{-1}) [coverage (%)]													
	Total				CR				SR					
Class	Stage IV	NSSL WRF	Difference	*P* value	Stage IV	NSSL WRF	Difference	*P* value	Stage IV	NSSL WRF	Difference	*P* value
1A	1.94 (22.11)	1.62 (29.40)	−16% (33%)	**0.061 (0.035)**	9.18 (2.96)	9.94 (3.32)	8% (13%)	0.367 (0.602)	1.05 (19.70)	0.74 (26.77)	−30% (36%)	**3×10^{-11} (0.028)**
1B	2.55 (17.00)	1.94 (25.81)	−24% (52%)	**2×10^{-5} (1×10^{-4})**	10.46 (3.85)	10.35 (3.80)	1% (1%)	0.730 (0.838)	1.11 (14.03)	0.76 (22.91)	−32% (63%)	**1×10^{-28} (4×10^{-5})**
1C	3.21 (12.46)	1.54 (20.82)	−52% (67%)	**1×10^{-13} (6×10^{-5})**	11.61 (4.14)	9.69 (3.04)	−17% (−27%)	**0.002 (0.012)**	1.12 (9.54)	0.66 (18.88)	−41% (98%)	**5×10^{-21} (3×10^{-6})**
2A	2.23 (16.00)	1.77 (18.33)	−21% (15%)	**0.032** (0.262)	9.80 (2.49)	10.46 (2.41)	7% (−3%)	**0.070** (0.755)	1.09 (13.91)	0.74 (16.35)	−32% (18%)	**9×10^{-16}** (0.203)
2B	2.60 (12.29)	2.08 (12.06)	−20% (−2%)	**0.003** (0.974)	10.75 (2.88)	10.44 (2.74)	3% (−5%)	0.891 (0.504)	1.13 (10.12)	0.79 (10.05)	−30% (−1%)	**1×10^{-32}** (0.830)
2C	3.12 (12.01)	2.30 (17.72)	−26% (48%)	**1×10^{-4} (9×10^{-4})**	11.28 (3.86)	10.94 (3.97)	−3% (3%)	0.657 (0.894)	1.16 (9.10)	0.77 (14.82)	−34% (63%)	**2×10^{-19} (2×10^{-4})**
Mean	2.61 (15.06)	1.91 (20.10)	−26% (33%)	**6×10^{-19} (4×10^{-8})**	10.54 (3.35)	10.33 (3.21)	−2% (−4%)	0.925 (0.340)	1.11 (12.50)	0.75 (17.70)	−32% (42%)	**1×10^{-112} (1×10^{-9})**

during the night. These results also provide a strong support of the higher amplitude diurnal cycles for these two classes observed in Fig. 5. The averaged CR and SR intensities for all six classes are 10.54 and 1.11 mm h^{-1}, respectively, which is consistent to the finding in Feng et al. (2012) where CR intensity is an order of magnitude higher than SR. Conversely, the averaged CR and SR coverages are 3.35% and 12.50% of the study domain, respectively, where the SR coverage is almost four times as large as the CR one.

Results for NSSL-WRF are shown in Figs. 7b and 7d, along with the mean values, biases from observations, and BEST *P* values between the two populations (Stage IV and NSSL-WRF) listed in Table 3. For total precipitation, the issue of lower intensity but broader coverage exists for each SOM class except for class 2B, which will be discussed later. This issue is even more prominent for SR, where larger negative biases in precipitation intensity and positive biases in precipitation coverage are found. For the CR component, deviations from observations still exist, but the overall magnitudes of intensity and coverage match better than total precipitation and SR portion. Through the BEST test between Stage IV observations and NSSL-WRF simulations, the overall *P* values for each class's CR intensity and coverage are all above the 0.05 significance level (except for class 1C featuring nocturnal upscale growth whose CR simulation significantly different from observation), indicating the hypothesis that both observations and simulations are from the same population cannot be rejected. Meanwhile, significant differences in the biases for SR component are proven for almost all classes except for class 2A and 2B's coverages. In summary, although the simulated precipitation amounts have good agreement with the observations as shown in Fig. 6, NSSL-WRF has weaker precipitation intensity but larger coverage, and this discrepancy is more prominent for the SR portion. In contrast, the CR intensity and coverage are better simulated for each SOM class.

As mentioned above, class 2B differs from the other classes by having a negative bias in precipitation coverage for total (−2%), CR (−5%), and SR (−1%), whereas the remaining classes of simulations have large positive biases in both the total and SR coverages. As a result, the undersimulated nocturnal precipitation amount (Fig. 5e) can be attributed to the missing precipitation coverage where the CR component (−5%) is more to blame than SR (−1%). Another interesting category is class 1C, which also has a lack of nocturnal precipitation (Fig. 5c) but with a larger deficit. Different from the class 2B's slight negative bias in CR coverage, this class significantly undersimulates both CR intensity (−17%) and coverage (−27%). Thus for these two

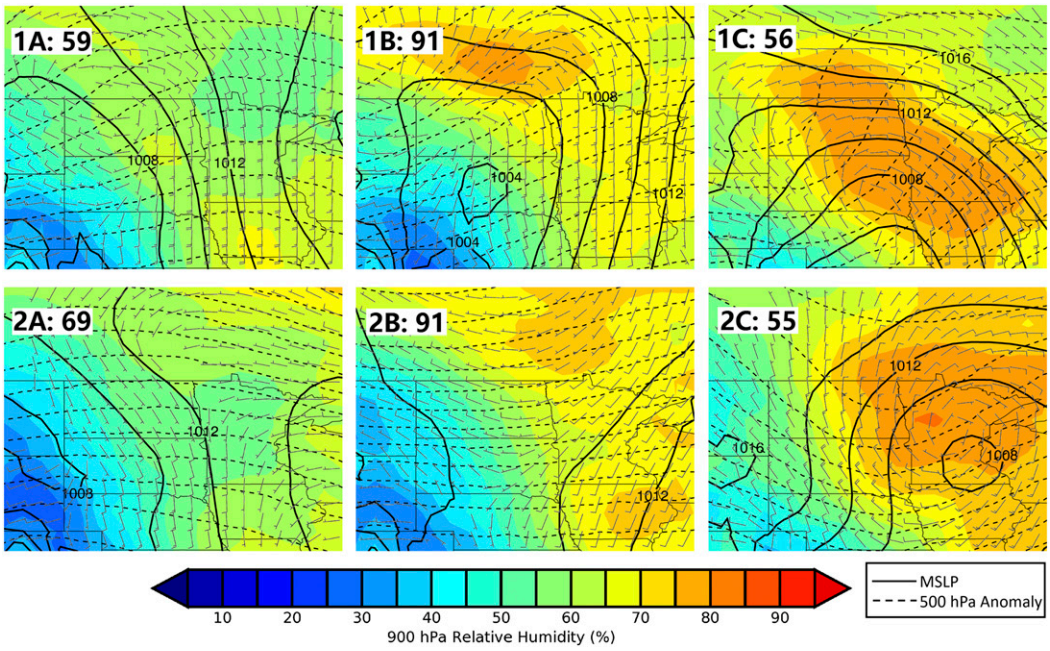

FIG. 8. As in Fig. 2, but for the NGP SOM.

classes, the undersimulated CR precipitation is the major error source for the deficit in nocturnal precipitation, but class 1C is more significant. The conclusion of missing CR for those two classes is in agreement with earlier discussions that NSSL-WRF simulation is incapable of capturing convective upscale growth based on Fig. 5c.

b. The Northern Great Plains

1) ANALYSIS OF SOM PATTERNS

Compared to the SGP, atmospheric patterns for the NGP SOM are more reliant on stronger synoptic-scale forcing with half of the classes (1B, 1C, 2C) featuring well defined midlatitude cyclones (Fig. 8). Given the location of this region near the geographic center of the continent, this should make sense; barring surface sources such as evapotranspiration or irrigation, the primary source for water vapor is large-scale advection associated with these systems. For the rest of NGP patterns, class 1A has stronger southerly flow at the surface underneath southwesterly flow at 500 hPa. Classes 2A and 2B have poorly defined surface features with slight ridging and near-zonal flow, respectively. Those patterns are more related to weaker synoptic-scale forcing.

Although patterns were classified independently for the two regions, it is worth noting that precipitation days could occur concurrently, and this is the case for 163 days (39% of the NGP cases). Not surprisingly, cases with well-defined midlatitude cyclones were most likely to be linked to each other. Nearly one-third of the overlapped cases (47) were linked to SGP class 1C. 38 of these 47 classes were associated with NGP classes 1B, 1C, and 2C. On the opposite end of the spectrum, cases with the least similarity included SGP class 1A (8 cases), and NGP 1A and 2A (18 cases for each).

Specific examples of precipitation cases are shown for each class within the NGP SOM in Fig. 9. Similar to the SGP, classes 1A–1C (top row of SOM) feature southwesterly flow at 500-hPa level with surface patterns ranging from southerly flow ahead of an approaching surface cyclone (class 1A), to the warm sector just northeast of a cyclone (class 1B), to the warm front north of the cyclone (class 1C). Class 2C can also be grouped with these types of cases as the domain of interest is located just northwest of the surface low underneath the upper-level trough, a typical pattern associated with Colorado lows. In contrast, subtropical high pressure centers in classes 2A–2B are located south of the study domain with precipitation occurring at the peripheries of the high pressure centers.

Similar to what was found for the SGP SOM, NGP patterns have preferred periods of occurrence (Fig. 10). Classes featuring the stronger midlatitude cyclones are most common from April–June. While class 1C is primarily confined to these months, classes 1B and 2C also have an increase in cases in August and September.

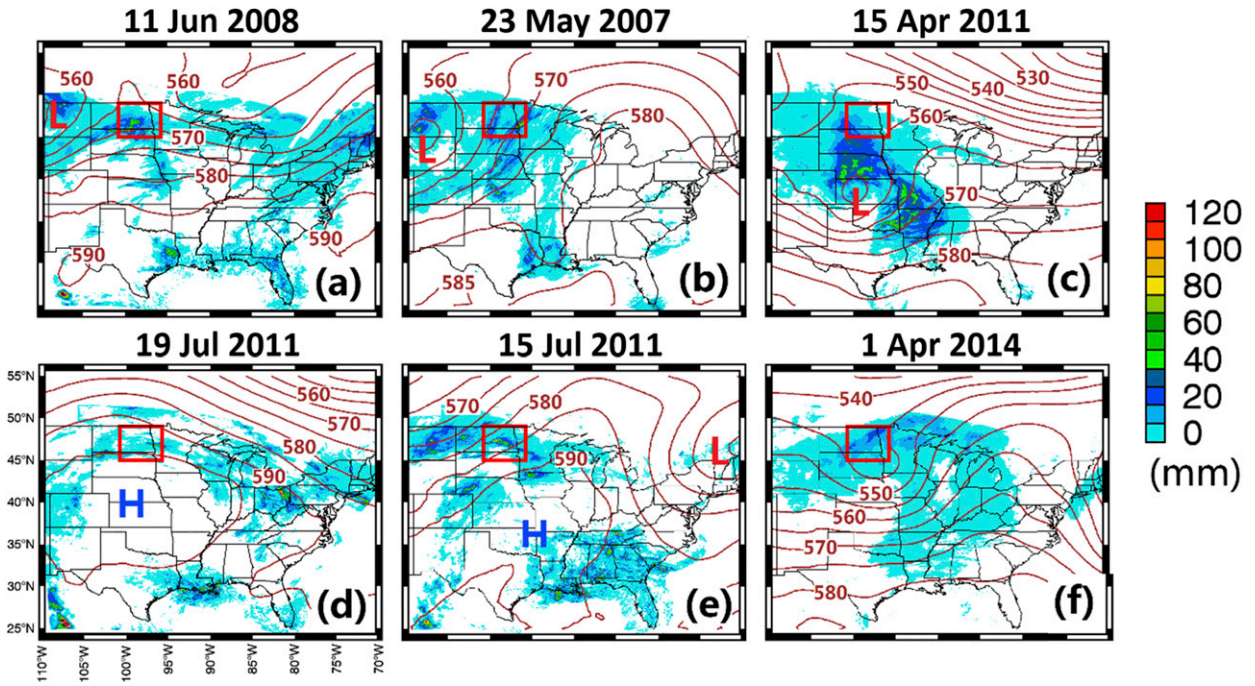

FIG. 9. Examples of precipitation cases for each class within the NGP SOM (a) class 1A: storm on 11 June 2008, (b) class 1B: 23 May 2007, (c) class 1C: 15 April 2011, (d) class 2A: 19 July 2011, (e) class 2B: 15 July 2011, and (f) class 2C: 1 April 2014.

Of the patterns with weaker surface features, classes 2A and 2B are predominately summer patterns with peak occurrence during July. Class 1A, which has southwesterly flow aloft over southerly near-surface winds, occurs throughout the season with only a minimal amount of cases in September. Whereas the SGP SOM showed top-down variability in seasonality, this feature is less apparent for NGP.

2) ANALYSIS OF THE DIURNAL CYCLE

Figure 11 shows the diurnal cycles of domain-averaged precipitation rates from Stage IV observations and

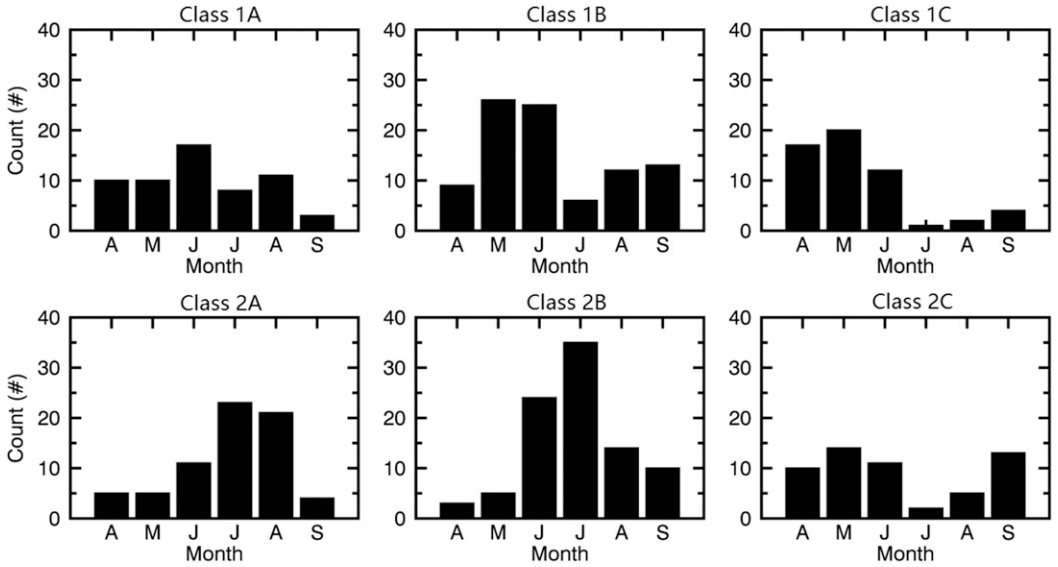

FIG. 10. As in Fig. 4, but for the NGP SOM.

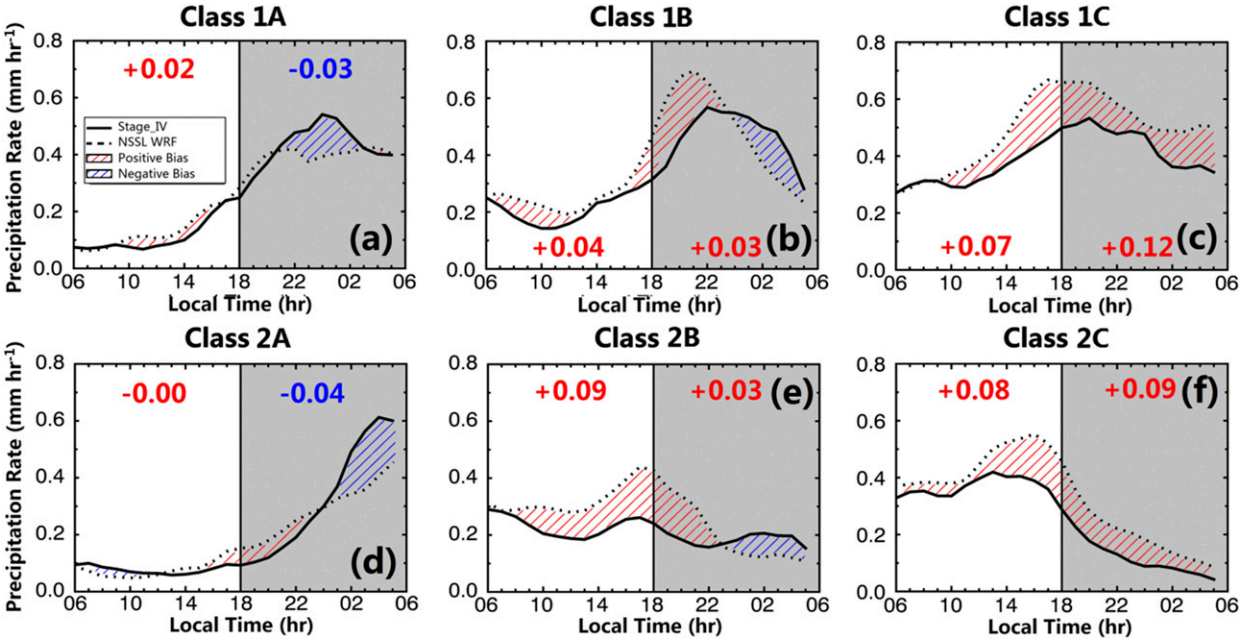

FIG. 11. As in Fig. 5, but for the NGP.

NSSL-WRF simulations for each class within the NGP SOM. Classes 1A, 1B, and 2A demonstrate different diurnal patterns compared to classes 1C, 2B, and 2C. The former is consistent with the typical GP diurnal pattern with nocturnal precipitation maxima, which resemble the SGP classes 1B, 1C, 2B, and 2C. The latter cases have less diurnal variability and their peaks occur around 1800 LT, which is similar to the SGP classes 1A and 2A (Figs. 5a and 5d), and the lack of day/night contrast may indicate the dominance of SR.

For the visualization of the day and night contrast, the NGP observed/simulated daytime/nighttime averaged precipitation amounts and biases for each SOM class are presented in Fig. 12 and listed in Table 1. From the perspective of observed precipitation, classes 1A, 1B, and 2A's nocturnal amounts are notably greater than the daytime. Nighttime precipitation is slightly higher (lower) than the daytime for the class 1C (2B), but class 2C receives much more precipitation during the day. The observed mean precipitation differences between day and night are further examined using BEST test as shown in Table 2. The *P* values among classes 1A, 1B, and 2A are relatively large, indicating they have similarity as demonstrated by the diurnal oscillation indicative of nocturnal upscale growth (Fig. 11). Conversely, classes 1C, 2B, and 2C are significantly different from the other classes. Comparing Fig. 12 to Fig. 6, one can easily notice there is less spatial variability in NGP than SGP, indicating the SR portion is more dominant over the NGP.

Strong positive biases for NSSL-WRF are found for SR dominated classes (1C and 2C, counterparts of SGP classes 1A and 2A) for both day and night. Classes 1A and 2A (counterparts of SGP classes 1C and 2B) correspond to the best daytime match (differences of 0.22 and −0.01 mm) and the largest negative nocturnal biases (−0.39 and −0.44 mm). Despite these similarities, regional differences still exist. For example, compared to the SGP classes 1A and 2A, NGP's SR dominated classes (1C and 2C) have notable diurnal variations (slightly increase and significant decrease from day to night, respectively), and there is almost no diurnal variation in class 2B whose SR component is not as pronounced. Note that class 2B is poorly defined by including distinct surface and upper-level features from the original 28-class SOMs, which cannot fully account for other variabilities in addition to the CR versus SR separation as explicit as the SGP.

3) ANALYSIS OF PRECIPITATION TYPE, INTENSITY, AND COVERAGE

To understand how precipitation is partitioned between CR and SR, box plots of precipitation are shown for the NGP SOM (Fig. 13). Stage IV observations demonstrate that classes 1C and 2C correspond to the lowest total/CR/SR intensity (Fig. 13a and Table 4) and the largest total/SR coverage (Fig. 13c), which confirms that SR is the primary precipitation component for these two classes. Since these two classes include the strongest midlatitude cyclones and a well-defined 500-hPa trough,

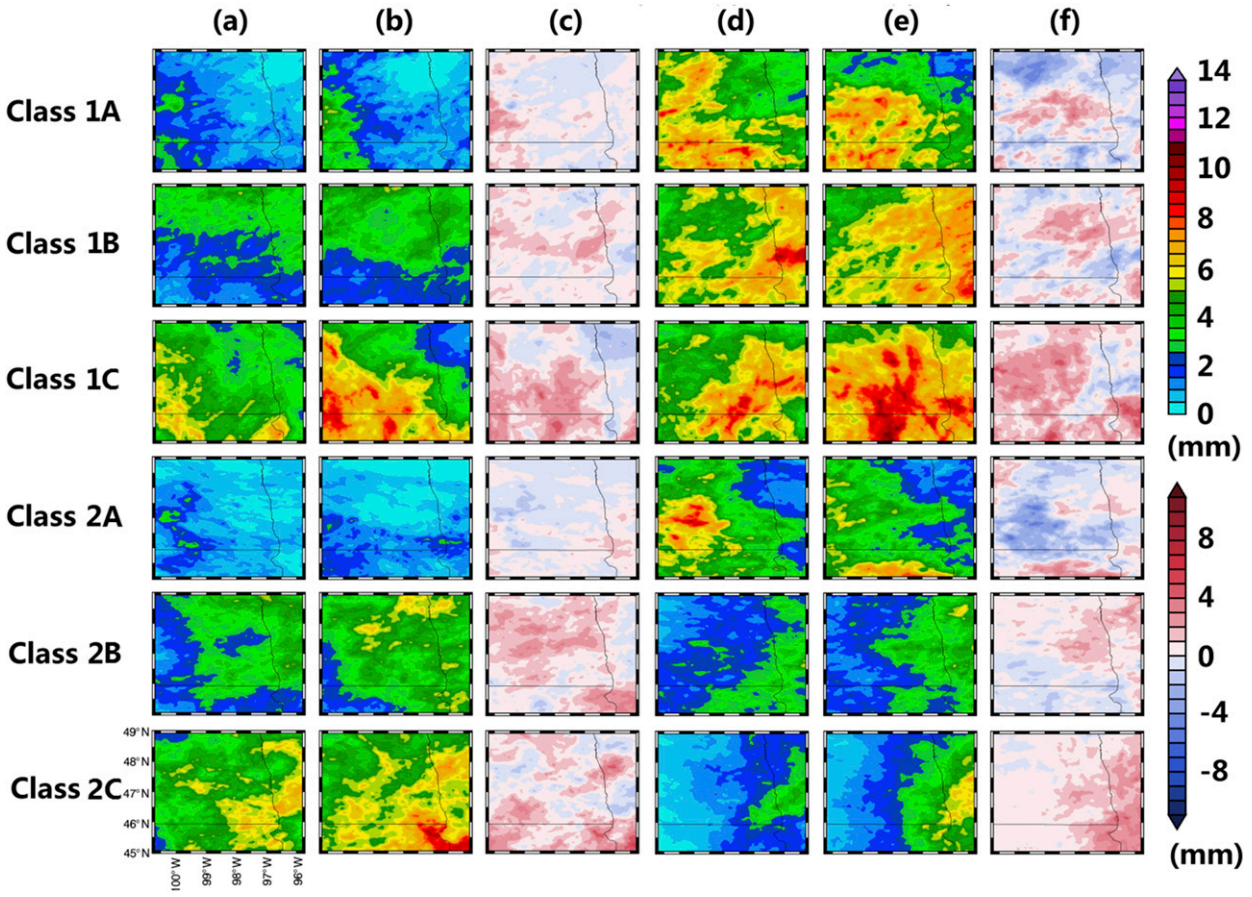

FIG. 12. As in Fig. 6, but for the NGP.

this makes meteorological sense as these environments support widespread and longer-lasting SR precipitation. Classes 1A, 1B, 2A, and 2B have higher overall intensity and their total and SR coverages are almost half of classes 1C and 2C (mean values are shown in Table 4), which suggests that convection is more important for these four classes. Comparing the mean observation values in Table 4 with those in Table 3, the NGP's total, CR, and SR intensities are all lower than the SGP, but NGP features 2.9% (4.06%) more total (SR) coverage and 0.88% less CR coverage than the SGP. From the perspective of NSSL-WRF simulations, NGP (Table 4) has similar patterns as the SGP (Table 3) for each of the SOM classes, which are 1) undersimulation in total and SR intensity but oversimulation in total and SR coverage, and 2) overall good match in CR intensity and coverage. However, the incorrect simulation of SR precipitation is a more prominent issue in the NGP than the SGP.

As revealed by Mahoney and Lackmann (2007), the north side of the CONUS is the upstream region of convection where the storms commonly initiate. After the southeastward propagation, they reach the mature stage in the downstream side in the south. Through the comparison with observations, NSSL-WRF demonstrates better overall performance over the SGP than the NGP. This regional difference may reflect the model's potential problem in simulating the storm at its different life stages (i.e., initiation vs mature). However, further evaluation involving storm tracking is needed to support this hypothesis.

c. *Evaluation of NSSL-WRF simulations using performance matrices*

Each region's 2D 24-h accumulated precipitation field is examined through the performance matrices defined previously. The mean values of these indices separated by SOM classes over the SGP and NGP regions are shown in Table 5. To facilitate comparisons, these statistical results are illustrated in Taylor diagrams in Fig. 14, where the correlation coefficient R and AI are shown as the rotation angle from the vertical axis, and NSTD and ratio are shown as the distance from the coordinate origin for the SGP (Figs. 14a and 14b) and the NGP (Figs. 14c and 14d).

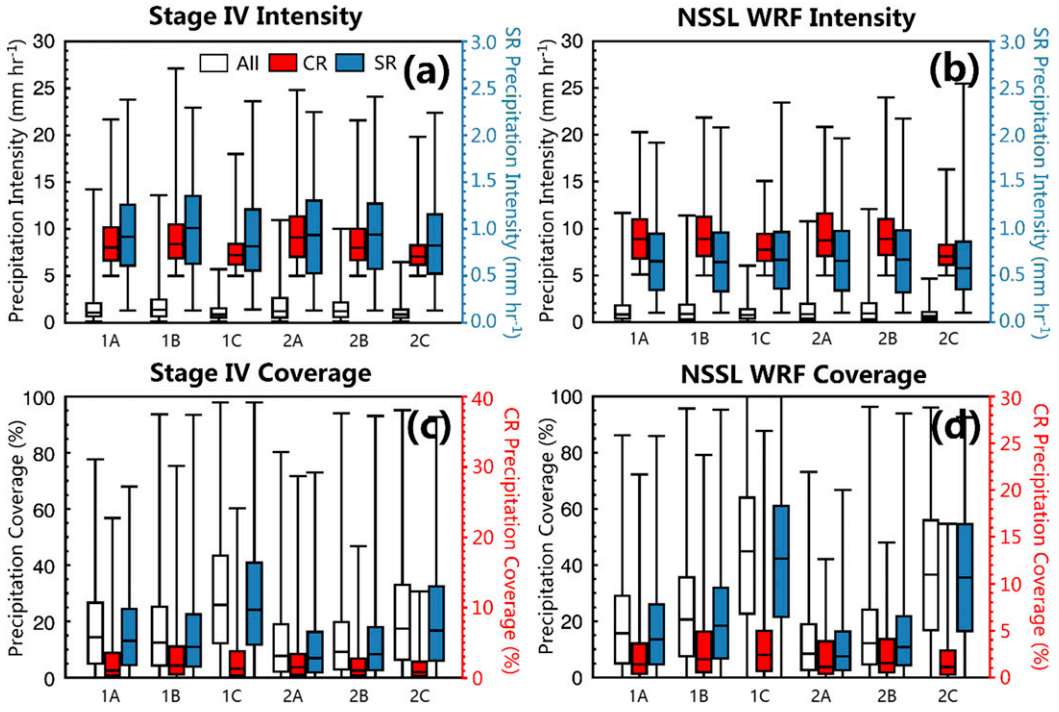

FIG. 13. As in Fig. 7, but for the NGP.

By comparing the SGP classes 1A–1C (southwesterly flow aloft), the best performance (highest R and AI) is found for SR dominant class 1A, which corresponds to the largest positive bias (highest ratio) and spatial variation (highest NSTD). Meanwhile, class 1C (CR with upscale growth) has the worst performance due to missing nocturnal convection, while class 1B falls between these two classes. These results are expected because both the R and AI scores emphasize the collocation between simulations and observations. Since class 1A has the largest all/SR coverage (Fig. 7, Table 3), collocation is reached more easily. Similar conclusions can also be drawn from the classes 2A–2C in the SGP SOM (northwesterly to zonal flow at 500 hPa), where the R and AI scores range from the highest in class 2A to lowest in class 2B. In general, the averaged ratio from six classes over the SGP region is 1.065, suggesting that the NSSL-WRF simulated precipitation is comparable to the Stage IV observation. In the NGP (Figs. 14c and 14d), the best performance is also found for SR dominated regimes (classes 1C and 2C). There are positive biases for all six classes with an average ratio of 1.23, which is 15.5% greater than the SGP average.

Separating performance by overall meteorological regime, Fig. 14 demonstrates that the extratropical cyclone impacted classes (SGP: 1A, 1B, 1C, 2B; NGP: 1A, 1B, 1C, and 2C) generally outperform their counterpart classes (SGP: 2A, 2C; NGP: 2B and 2C) under the influence of the subtropical ridge for both regions. Through the synthesis of all cases within different synoptic schemes (extratropical cycle vs subtropical ridge), the former has higher R and AI values than the latter (SGP: 0.266, 0.486 vs 0.209, 0.458; NGP: 0.258, 0.479 vs 0.226, 0.445).

4. Summary and discussion

SOMs were used to classify atmospheric states for the upper 90% of precipitating days over the SGP and NGP domains. With no a priori knowledge of the number of meteorological patterns responsible for precipitation, this methodology classified patterns to 28-class (7×4) SOMs. With the premise that pattern recognition can be used to identify expected WRF precipitation biases, these larger SOMs were averaged to final 6-class maps. Region specific patterns were identified, and regime-dependent biases were found within NSSL-WRF from 2007 to 2014.

Following the SOM classes generated, precipitation events over both the SGP and NGP regions were examined from two perspectives: the primary precipitation type (CR vs SR) and the dominant synoptic pattern (extratropical cyclone vs subtropical ridge). From Stage IV observations, distinct characteristics are found regarding the diurnal cycle and precipitation intensity/coverage. Some of those features can be well simulated

TABLE 4. As in Table 3, but for the NGP SOM.

	Precipitation Intensity (mm h^{-1}) [coverage (%)]											
	Total				CR				SR			
Class	Stage IV	NSSL WRF	Difference	*P* value	Stage IV	NSSL WRF	Difference	*P* value	Stage IV	NSSL WRF	Difference	*P* value
1A	1.59 (17.83)	1.40 (19.06)	−12% (7%)	0.421 (0.485)	8.82 (2.56)	9.32 (2.69)	6% (5%)	0.554 (0.866)	0.96 (16.35)	0.67 (17.53)	−31% (7%)	**3×10^{-11}** (0.495)
1B	1.80 (16.85)	1.45 (24.05)	−19% (43%)	**0.008 (3×10^{-5})**	9.04 (3.16)	9.43 (3.50)	4% (11%)	**0.079** (0.503)	1.00 (14.88)	0.67 (21.90)	−33% (50%)	**6×10^{-27} (2×10^{-5})**
1C	1.20 (29.50)	1.02 (44.22)	−15% (50%)	**0.039 (3×10^{-6})**	7.52 (2.89)	8.14 (3.72)	8% (29%)	0.709 (0.300)	0.90 (28.08)	0.69 (42.39)	−23% (51%)	**1×10^{-5} (5×10^{-6})**
2A	1.77 (12.56)	1.47 (13.08)	−17% (4%)	**0.097** (0.634)	9.61 (2.43)	9.69 (2.33)	1% (4%)	0.567 (0.606)	0.94 (11.18)	0.67 (11.82)	−29% (6%)	**8×10^{-14}** (0.578)
2B	1.61 (13.02)	1.47 (16.50)	−9% (27%)	**0.827 (0.003)**	8.68 (2.03)	9.48 (2.75)	9% (35%)	**0.019 (0.021)**	0.93 (11.80)	0.68 (14.86)	−27% (26%)	**9×10^{-16} (0.005)**
2C	1.09 (22.30)	0.82 (37.85)	−25% (70%)	**0.001 (5×10^{-6})**	7.43 (1.74)	7.48 (2.21)	1% (27%)	0.953 (0.339)	0.86 (21.50)	0.65 (36.82)	−24% (71%)	**3×10^{-7} (4×10^{-6})**
Mean	1.54 (17.96)	1.30 (24.79)	−16% (38%)	**0.001 (4×10^{-6})**	8.57 (2.47)	9.00 (2.89)	5% (17%)	0.953 (0.339)	0.94 (16.56)	0.67 (23.17)	−29% (40%)	**3×10^{-7} (4×10^{-6})**

by NSSL-WRF, while others have more significant disagreements. These results can be briefly summarized as follows.

1) For the SGP, synoptic patterns can be divided into states with surface warm fronts (classes 1A, 1B), lying within the warm sector of an extratropical cyclone (classes 1C, 2B), and more quiescent states under northwesterly flow aloft (classes 2A, 2C). Drylines within the vicinity of the analysis region were common and the intensity of the dryline varied across the SGP SOM. Unsurprisingly, patterns demonstrated preferred months of occurrence with more active classes more common during the spring/early summer and fall. More quiescent states with northwest flow aloft were more common during the summer as the jet stream lifted north of this region. Of the six SOM classes, four classes with more convective activity have lower daytime but higher nocturnal precipitation amount. Moreover, the largest day/night contrast is found in classes 1C and 2B, indicating notable upscale growth in the night. These results provide a strong support to the well-documented GP nocturnal maximum precipitation diurnal pattern. However, this conclusion only works for the classes with more convection, and is not valid for the SR dominated classes (1A and 2A, taking up 30% of all heavy precipitation events).
2) Patterns for the NGP were more focused on stronger baroclinic wave activity, with half of the states featuring prominent extratropical cyclones. These states (classes 1A, 1B, 1C, 2C) were most common during the spring/early summer as seen for the SGP SOM. Two states were more common during the summer months (classes 2A, 2B), but featured stronger near-zonal flow aloft compared to the summer states seen for the SGP. For the NGP, half of the classes have the typical GP diurnal pattern with nocturnal precipitation maxima while the other half have less diurnal variability with peaks occurring around 1800 LT. In combination with the conclusion drawn for the SGP, correct simulation of the nocturnal maximum precipitation remains a challenge in NSSL-WRF. From the perspective of regional difference (Tables 3 and 4), higher precipitation intensity and larger precipitation amount is observed over the SGP than the NGP.
3) Comparing the NSSL-WRF simulations with Stage IV observations from the perspective of the diurnal cycle, both regions' SR dominated classes are well simulated by NSSL-WRF, while classes with nocturnal upscale growth correspond to the largest negative bias at night. This is in agreement with prior studies that NSSL-WRF

TABLE 5. Performance indices of the NSSL-WRF simulation for classes within the SGP and the NGP SOMs.

	Indices	Class 1A	Class 1B	Class 1C	Class 2A	Class 2B	Class 2C
SGP	CC	0.294	0.259	0.249	0.249	0.176	0.224
	NSTD	1.180	1.052	0.769	1.057	0.961	1.059
	AI	0.496	0.488	0.473	0.473	0.441	0.470
	Ratio	1.296	1.153	0.776	1.069	0.914	1.187
NGP	CC	0.273	0.201	0.334	0.164	0.198	0.352
	NSTD	1.079	1.153	1.217	1.047	1.341	1.248
	AI	0.485	0.451	0.518	0.399	0.426	0.533
	Ratio	1.042	1.178	1.347	1.063	1.355	1.398

ends convection too soon (Clark et al. 2011). Transitional classes connecting SR to CR dominated cases match the observations in overall magnitude, but instead of missing nocturnal convection, simulated precipitation peaks too early, causing a daytime positive bias and nocturnal negative bias. Simulated CR intensity and coverage have good agreement with the observations for both regions, but NSSL-WRF undersimulates SR precipitation intensity and oversimulates SR coverage.

4) For different SOM classes, precipitation simulated by NSSL-WRF demonstrates distinct behaviors (intensity, coverage, CR vs SR partitioning, diurnal cycle) compared to the Stage IV observation. Through the synthesis of all cases within different synoptic schemes (extratropical cycle vs subtropical ridge) over both regions, the NSSL-WRF has better performance for the former than the latter. By separating the classes using dominant precipitation type,

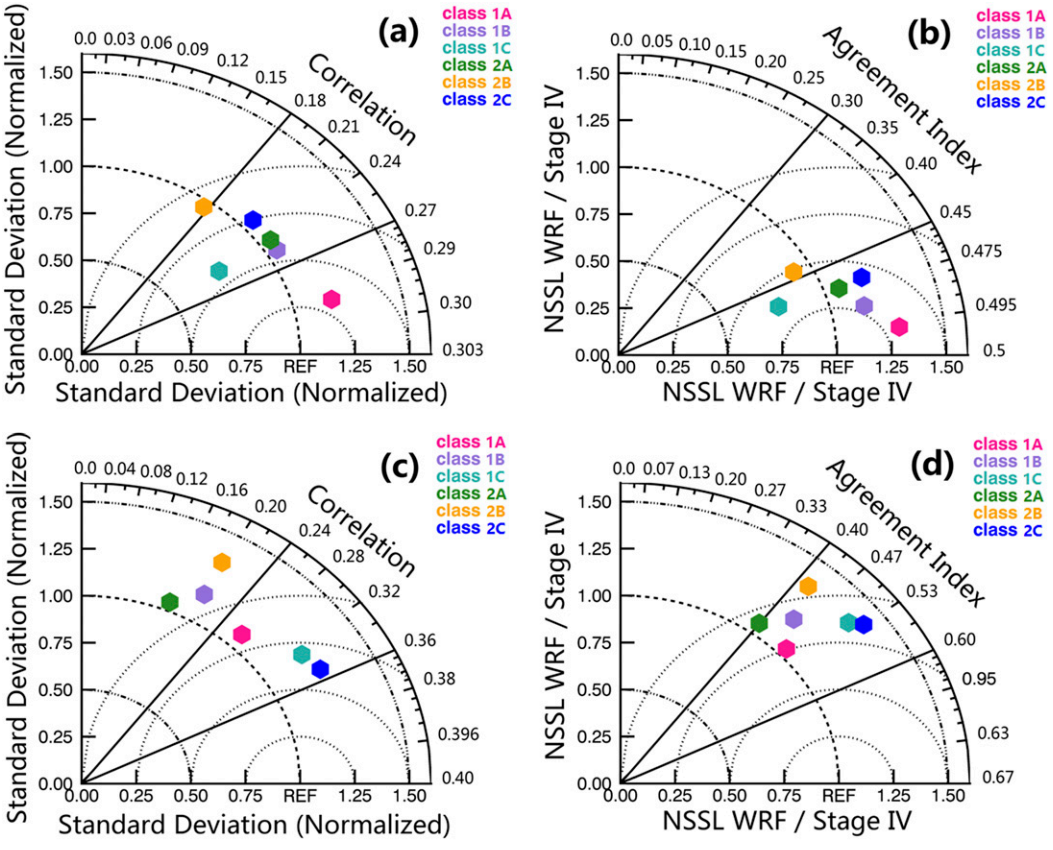

FIG. 14. Taylor diagrams for (a),(c) normalized standard deviation vs correlation and (b),(d) ratio vs agreement index for each class in the (top) SGP and (bottom) NGP SOMs. In the polar coordinate system, the distance from origin to a certain point represents the normalized standard deviation or ratio, and the angle from the x direction represents the correlation or agreement index between NSSL-WRF and Stage IV observation.

the persistent better performance is strongly tied to SR dominated cases, whereas the missing nocturnal precipitation remains the major issue for the simulation of convection. The demonstrated utility of SOMs provides the new insights about the model evaluation under different atmospheric contexts.

This study has only scraped the surface of what is possible with regime-based assessment of CAM output. We also note the RRC-based separation of CR versus SR has intrinsic limitation to account for the differences in heating profiles, spatial distribution, environmental factors, etc. between those two storm portions. Provided that an adequate database of CAM simulations is available, SOMs (and other automated pattern classification algorithms) show promise to be a useful tool for users of CAM forecasters. Outstanding questions include applying these concepts to model forecasts (e.g., position errors in simulated features such as the dryline), to analyzing proxies for hazardous weather (e.g., updraft helicity), and extending this work to probabilistic studies that include CAM ensembles. Ideally, these types of algorithms could be run in real-time to provide forecasters information on expected biases as model simulations are completed.

Acknowledgments. This research was supported by the Climate Model Development and Validation (CMDV) program funded by the Office of Biological and Environmental Research in the U.S. Department of Energy Office of Science under Grant DE-SC0017015 at the University of Arizona, and by the NOAA R2O project with Award NA15NWS468004 at the University of North Dakota and subcontracted to the University of Arizona. Work at the Pacific Northwest National Laboratory (PNNL) was supported by the Climate Model Development and Validation program and the Water Cycle and Climate Extreme Modeling science focus area funded by the Office of Biological and Environmental Research in the U.S. Department of Energy Office of Science. PNNL is operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract DE-AC05-76RL01830. The authors thank Mr. Alex Trellinger for assistance with the drafting of several of the figures and Dr. Matthew Gilmore and three anonymous reviewers for providing excellent feedback on drafts of this manuscript. NSSL-WRF simulations were provided by Drs. Adam Clark, Jack Kain, and Patrick Marsh. NARR data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, from their website at http://www.esrl.noaa.gov/psd/. The 4-km gridded Stage IV data were downloaded from National Center for Atmospheric Research (NCAR)/University Corporation for Atmospheric Research (UCAR) Earth Observing Laboratory (EOL) using the link http://data.eol.ucar.edu/dataset/21.093 accessed on 20 November 2016. The statistical data generated by this study can be obtained from Dr. Xiquan Dong (xdong@email.arizona.edu).

APPENDIX

A Brief Review and the Configuration of SOM Method Used in This Study

As one of the neural network cluster techniques, SOMs offer advantages over methods such as principal component analysis (PCA) or empirical orthogonal functions (EOFs). Hewitson and Crane (2002) highlighted the fact that SOMs make no a priori assumptions about the data. This study also demonstrated that SOMs capture nonlinear characteristics of the data and represent more subtle variations in the data by placing more nodes in these regions. Liu et al. (2006) established SOMs, unlike EOFs, could classify patterns associated with a linear propagating sine wave. Further, the SOM was able to replicate these patterns with noise added. This feature was also noted in Reusch et al. (2007) who showed SOMs were more robust by allowing better identification of pattern mixing (e.g., physically unrealistic scenarios) versus PCA. In fact, they demonstrated that smaller SOMs shared many characteristics with PCAs.

Despite the advantages, SOMs are not a panacea for clustering analysis. As noted in the SOM review articles such as Sheridan and Lee (2011) and Liu and Weisberg (2011), users must make a number of decisions that can determine the success of their SOM. Besides selecting an initial number of nodes, there are a number of tunable parameters that complicates the initial creation of the SOMs. Despite this problem, studies such as Liu et al. (2006) and Kennedy et al. (2016) provide suggestions on appropriate settings for the creation of SOMs. For a basic example of a SOM, readers are encouraged to read Hewitson and Crane (2002) who used the technique to cluster MSLP patterns over the Northeast United States.

Since the reviews in Sheridan and Lee (2011) and Liu and Weisberg (2011), SOMs have been widely used to investigate a plethora of mesoscale phenomena and modeling systems. One such example is Nigro et al. (2017) who typed patterns over the Ross Ice Shelf to evaluate the Antarctic Mesoscale Prediction System (AMPS). While this study classified 2D MSLP patterns, SOMs have also been applied to 1D training sets such as

TABLE A1. SOM settings used in this study.

SOM settings	Value
Trials	10
Training length (stage 1)	No. of cases
Training length (stage 2)	No. of cases × 100
Learning rate (stage 1)	0.05
Learning rate (stage 2)	0.01
Neighborhood radius (stage 1)	Xdim-1
Neighborhood radius (stage 2)	1

vertical thermodynamic sounding profiles in this region (Cassano et al. 2016; Nigro et al. 2017). At lower-latitude locations, SOMs have been applied to soundings in near-storm environments (Nowotarski and Jensen 2013; Nowotarski and Jones 2018) or to characterize boundary layer wind properties in an alpine valley (Katurji et al. 2015). They have also been applied to gridded mesoscale model data to investigate parameters associated with tornado events (Anderson-Frey et al. 2017) or surface fronts (Hope et al. 2014). In general, SOMs have provided useful insight into model behavior and have been advocated for as a tool to identify model and forecast issues (Kolczynski and Hacker 2014).

In this study, the SOMs were generated using the freely available SOM_PAK software (Kohonen et al. 1996) and largely followed the methodology of Kennedy et al. (2016). Each SOM was randomly initialized 10 times using the ''vfind'' command, and the map with the lowest quantization error was saved. Settings for vfind including training length, learning rates, and neighborhood radii are summarized in Table A1. For additional details of this process, the reader is referred to Kennedy et al. (2016).

Like many objective classification schemes, a critical decision in the methodology is the number of classes chosen for the SOM feature map. If too few classes are chosen, important differences in atmospheric states may be smoothed out; too many and there are an inadequate number of precipitation cases within each state limiting the statistical significance of the results. With no a priori knowledge of the number of patterns associated with precipitation events, experience with prior work with SOMs, and the number of precipitation cases for each region (~400), larger 7 × 4 (28-class) SOMs were first created (Figs. A1–A4). Analysis of these initial SOMs revealed gradients in precipitation across the feature maps, but with significant class-to-class variability. To increase the robustness of statistics, and create a reasonable set of patterns forecasters could use for pattern recognition purposes, the 28-class SOMs were averaged to final 3 × 2 (6-class) SOMs.

The averaging process was determined objectively by comparing the Euclidean distance (similarity) of each pattern to the remainder of the patterns within the SOM (Fig. A5). As is typical of SOMs, more dissimilar

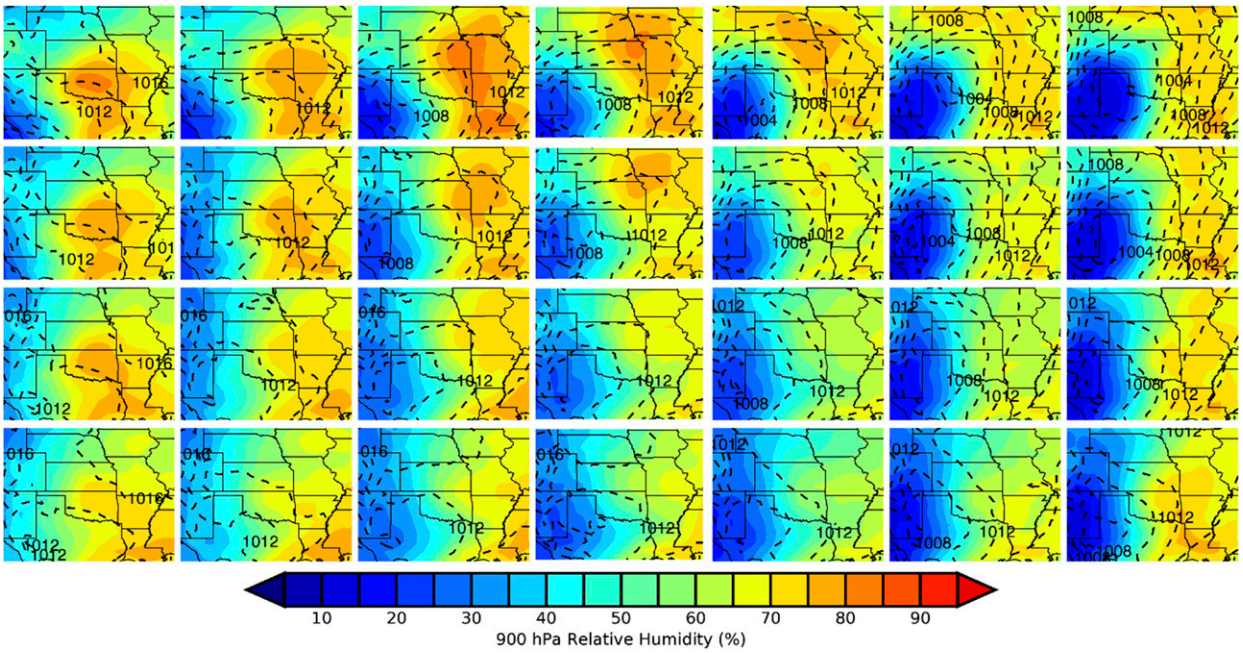

FIG. A1. Near-surface analyses for the 28-class (7 × 4) SGP SOM. MSLP is contoured with dashed lines while filled contours represent 900-hPa RH. Cool (warm) colors represent drier (moister) air.

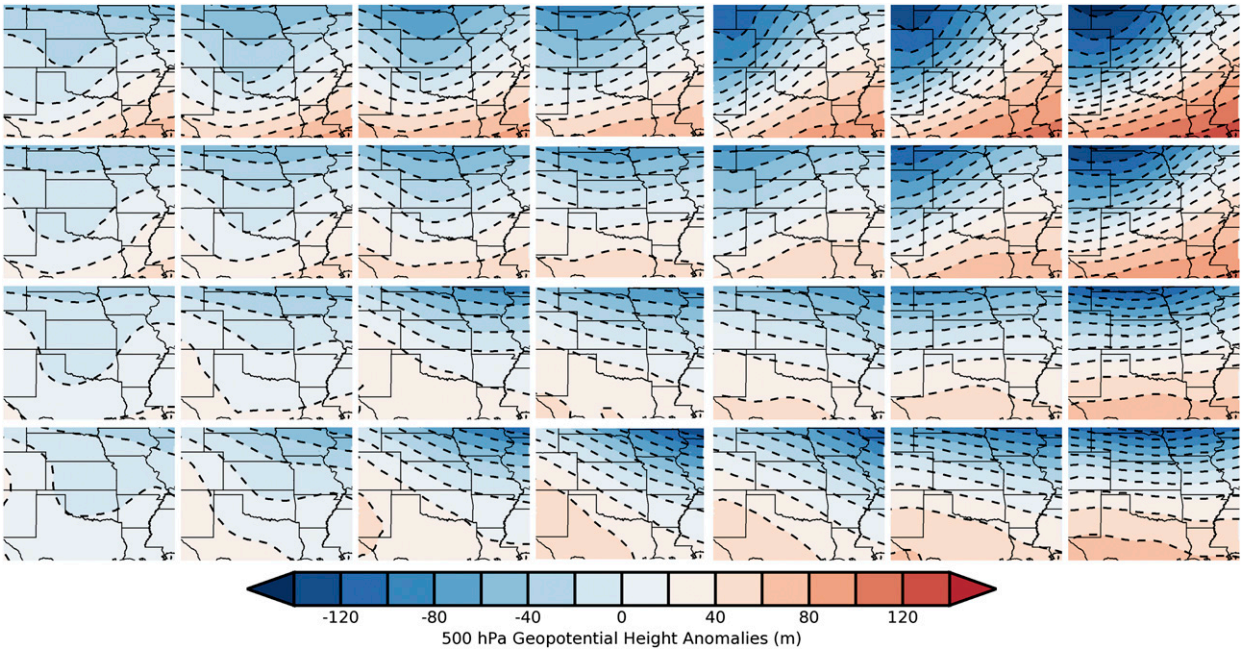

FIG. A2. The 500-hPa geopotential height anomalies for the 28-class (7 × 4) SGP SOM. Blues (reds) indicate negative (positive) height anomalies.

patterns were found in the corner of the SOMs. This led to an averaging process that included four patterns for the corners of the SOMs, and six patterns for middle patterns (Fig. A5). This two-staged methodology may seem overly complex compared to a simple one-step process of producing a 6-class SOM. Given the lack of a priori knowledge of how many ''typical'' synoptic events are associated with rainfall over the regions, the larger

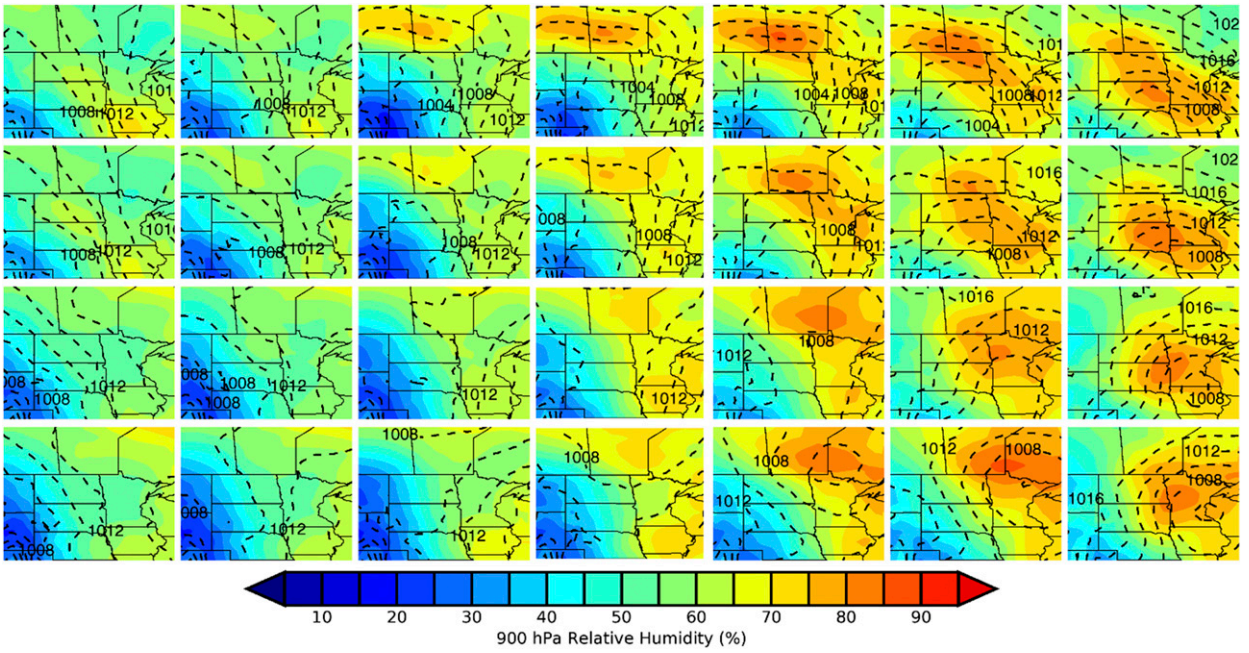

FIG. A3. As in Fig.A1, but for the 28-class (7 × 4) NGP SOM.

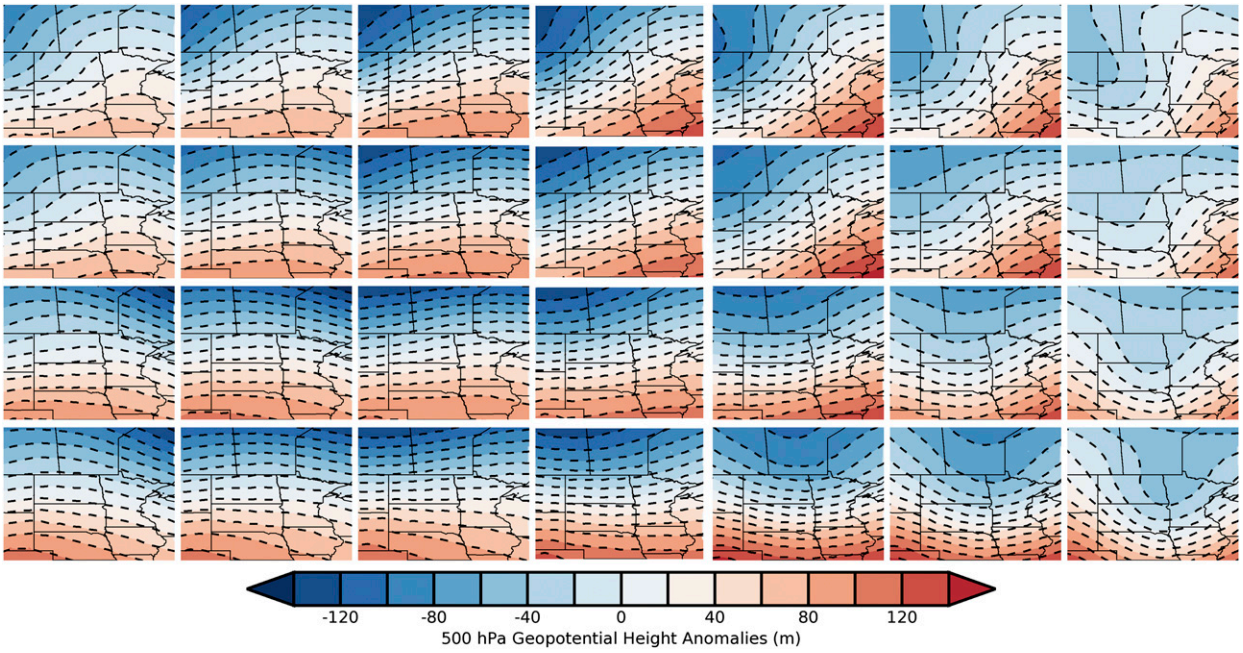

FIG. A4. As in Fig. A2, but for the 28-class (7 × 4) NGP SOM.

SOMs allowed for a better understanding of how representative final patterns were. Further, due to the nature of the SOM that places less similar cases more distant from each other within the feature map, the final 6-class SOMs had stronger gradients within the meteorological fields (e.g., MSLP). This is owed to averaging a fewer number of cases within the corner classes.

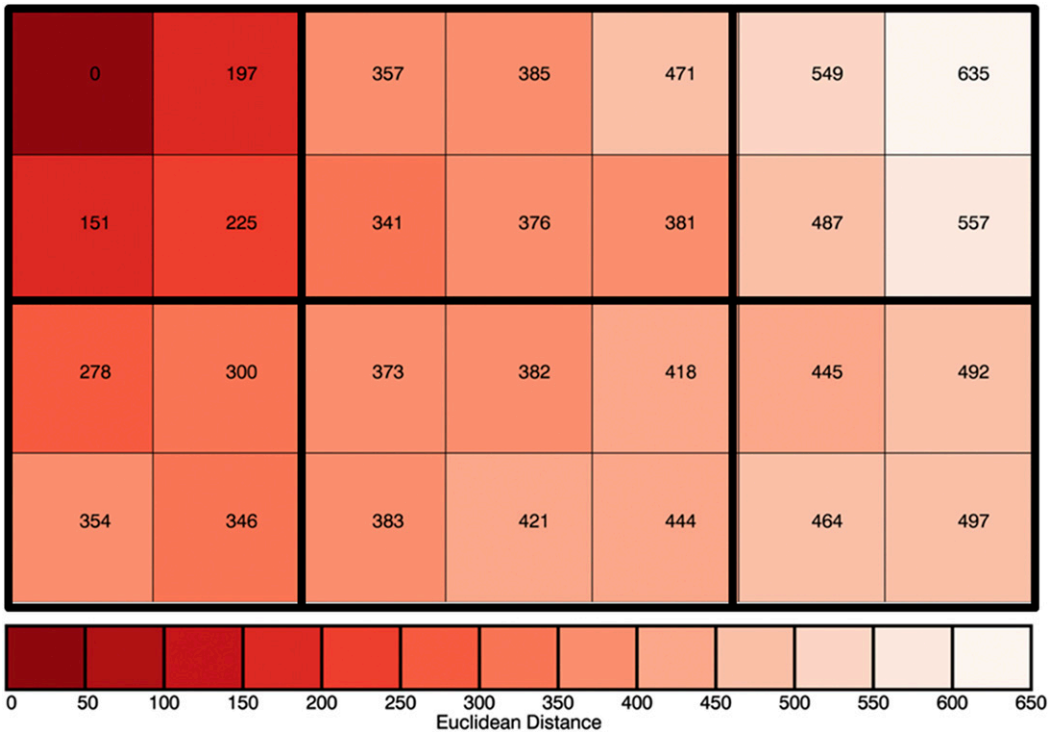

FIG. A5. Euclidian distance of each class from the upper-left class for the 28-class SGP SOM shown in Figs. A1 and A2. Dark lines represent boundaries for the classes averaged together for the final 6-class SOM.

REFERENCES

Anderson-Frey, A. K., Y. P. Richardson, A. R. Dean, R. L. Thompson, and B. T. Smith, 2017: Self-organizing maps for the investigation of tornadic near-storm environments. *Wea. Forecasting*, **32**, 1467–1475, https://doi.org/10.1175/WAF-D-17-0034.1.

Ashley, W. S., T. L. Mote, P. G. Dixon, S. L. Trotter, E. J. Powell, J. D. Durkee, and A. J. Grundstein, 2003: Distribution of mesoscale convective complex rainfall in the United States. *Mon. Wea. Rev.*, **131**, 3003–3017, https://doi.org/10.1175/1520-0493(2003)131<3003:DOMCCR>2.0.CO;2.

Baldwin, M. E., and K. E. Mitchell, 1997: The NCEP hourly multisensor U.S. precipitation analysis for operations and GCIP research. Preprints, *13th Conf. on Hydrology*, Long Beach, CA, Amer. Meteor. Soc., 54–55.

——, S. Lakshmivarahan, and J. S. Kain, 2001: Verification of mesoscale features in NWP models. Preprints, *Ninth Conf. on Mesoscale Processes*, Fort Lauderdale, FL, Amer. Meteor. Soc., 255–258.

Beard, K. V., D. B. Johnson, and D. Baumgardner, 1986: Aircraft observations of large raindrops in warm, shallow, convective clouds. *Geophys. Res. Lett.*, **13**, 991–994, https://doi.org/10.1029/GL013i010p00991.

Behrangi, A., K. Hsu, B. Imam, S. Sorooshian, G. J. Huffman, and R. J. Kuligowski, 2009: PERSIANN-MSA: A precipitation estimation method from satellite-based multispectral analysis. *J. Hydrometeor.*, **10**, 1414–1429, https://doi.org/10.1175/2009JHM1139.1.

Cassano, J. J., M. A. Nigro, and M. A. Lazzara, 2016: Characteristics of the near-surface atmosphere over the Ross Ice Shelf, Antarctica. *J. Geophys. Res. Atmos.*, **121**, 3339–3362, https://doi.org/10.1002/2015JD024383.

Chen, T., and J. A. Kpaeyeh, 1993: The synoptic-scale environment associated with the low-level jet of the Great Plains. *Mon. Wea. Rev.*, **121**, 416–420, https://doi.org/10.1175/1520-0493(1993)121<0416:TSSEAW>2.0.CO;2.

Clark, A. J., W. A. Gallus, and T. Chen, 2007: Comparison of the diurnal precipitation cycle in convection-resolving and non-convection-resolving mesoscale models. *Mon. Wea. Rev.*, **135**, 3456–3473, https://doi.org/10.1175/MWR3467.1.

——, and Coauthors, 2011: Probabilistic precipitation forecast skill as a function of ensemble size and spatial scale in a convection-allowing ensemble. *Mon. Wea. Rev.*, **139**, 1410–1418, https://doi.org/10.1175/2010MWR3624.1.

——, R. G. Bullock, T. L. Jensen, M. Xue, and F. Kong, 2014: Application of object-based time-domain diagnostics for tracking precipitation systems in convection-allowing models. *Wea. Forecasting*, **29**, 517–542, https://doi.org/10.1175/WAF-D-13-00098.1.

——, and Coauthors, 2018: The Community Leveraged Unified Ensemble (CLUE) in the 2016 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Bull. Amer. Meteor. Soc.*, **99**, 1433–1448, https://doi.org/10.1175/BAMS-D-16-0309.1.

Coffer, B. E., L. C. Maudlin, P. G. Veals, and A. J. Clark, 2013: Dryline position errors in experimental convection-allowing NSSL-WRF Model forecasts and the operational NAM. *Wea. Forecasting*, **28**, 746–761, https://doi.org/10.1175/WAF-D-12-00092.1.

Colle, B. A., and C. F. Mass, 2000: The 5–9 February 1996 flooding event over the Pacific Northwest: Sensitivity studies and evaluation of the MM5 precipitation forecasts. *Mon. Wea. Rev.*, **128**, 593–617, https://doi.org/10.1175/1520-0493(2000)128<0593:TFFEOT>2.0.CO;2.

Colman, B. R., 1990a: Thunderstorms above frontal surfaces in environments without positive CAPE. Part I: A climatology. *Mon. Wea. Rev.*, **118**, 1103–1122, https://doi.org/10.1175/1520-0493(1990)118<1103:TAFSIE>2.0.CO;2.

——, 1990b: Thunderstorms above frontal surfaces in environments without positive CAPE. Part II: Organization and instability mechanisms. *Mon. Wea. Rev.*, **118**, 1123–1144, https://doi.org/10.1175/1520-0493(1990)118<1123:TAFSIE>2.0.CO;2.

Cui, W., X. Dong, B. Xi, J. Fan, J. Tian, J. Wang, and T. McHardy, 2019: Understanding ice cloud-precipitation properties of three modes of mesoscale convective systems during PECAN. *J. Geophys. Res. Atmos.*, **124**, 4121–4140, https://doi.org/10.1029/2019JD030330.

Dai, A., 2001: Global precipitation and thunderstorm frequencies. Part II: Diurnal variations. *J. Climate*, **14**, 1112–1128, https://doi.org/10.1175/1520-0442(2001)014<1112:GPATFP>2.0.CO;2.

Davis, C., B. Brown, and R. Bullock, 2006: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, https://doi.org/10.1175/MWR3145.1.

Dey, S. R. A., N. M. Roberts, R. S. Plant, and S. Migliorini, 2016: A new method for the characterization and verification of local spatial predictability for convective-scale ensembles. *Quart. J. Roy. Meteor. Soc.*, **142**, 1982–1996, https://doi.org/10.1002/qj.2792.

Done, J., C. A. Davis, and M. Weisman, 2004: The next generation of NWP: Explicit forecasts of convection using the Weather Research and Forecasting (WRF) model. *Atmos. Sci. Lett.*, **5**, 110–117, https://doi.org/10.1002/asl.72.

——, G. Holland, C. Bruyère, L. Leung, and A. Suzuki-Parker, 2015: Modeling high-impact weather and climate: Lessons from a tropical cyclone perspective. *Climatic Change*, **129**, 381–395, https://doi.org/10.1007/s10584-013-0954-6.

Dudhia, J., 1989: Numerical observed study of convection during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Feng, Z., X. Dong, B. Xi, C. Schumacher, P. Minnis, and M. Khaiyer, 2011: Top-of-atmosphere radiation budget of convective core/stratiform rain and anvil clouds from deep convective systems. *J. Geophys. Res.*, **116**, D23202, https://doi.org/10.1029/2011JD016140.

——, ——, ——, S. A. McFarlane, A. Kennedy, B. Lin, and P. Minnis, 2012: Life cycle of midlatitude deep convective systems in a Lagrangian framework. *J. Geophys. Res.*, **117**, D23201, https://doi.org/10.1029/2011JD017012.

——, L. R. Leung, R. A. Houze Jr., S. Hagos, J. Hardin, Q. Yang, B. Han, and J. Fan, 2018: Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the United States. *J. Adv. Model. Earth Syst.*, **10**, 1470–1494, https://doi.org/10.1029/2018MS001305.

Fritsch, J. M., and R. E. Carbone, 2004: Improving quantitative precipitation forecasts in the warm season: A USWRP research and development strategy. *Bull. Amer. Meteor. Soc.*, **85**, 955–966, https://doi.org/10.1175/BAMS-85-7-955.

Galarneau, T. J., and L. F. Bosart, 2006: Ridge Rollers: Mesoscale disturbances on the periphery of cutoff anticyclones. *21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction*, Atlanta, GA, Amer. Meteor. Soc., 3.2, https://ams.confex.com/ams/WAFNWP34BC/techprogram/paper_94414.htm.

Gallus, W. A., 2002: Impact of verification grid-box size on warm-season QPF skill measures. *Wea. Forecasting*, **17**, 1296–1302, https://doi.org/10.1175/1520-0434(2002)017<1296:IOVGBS>2.0.CO;2.

Giangrande, S. E., S. Collis, A. K. Theisen, and A. Tokay, 2014: Precipitation estimation from the ARM distributed radar network during the MC3E campaign. *J. Appl. Meteor. Climatol.*, **53**, 2130–2147, https://doi.org/10.1175/JAMC-D-13-0321.1.

Goines, D. C., and A. D. Kennedy, 2018: Precipitation from a multiyear database of convection-allowing WRF simulations. *J. Geophys. Res. Atmos.*, **123**, 2424–2453, https://doi.org/10.1002/2016JD026068.

Herman, G. R., and R. S. Schumacher, 2016: Extreme precipitation in models: An evaluation. *Wea. Forecasting*, **31**, 1853–1879, https://doi.org/10.1175/WAF-D-16-0093.1.

Hewitson, B. C., and R. G. Crane, 2002: Self-organizing maps: Applications to synoptic climatology. *Climate Res.*, **22**, 13–26, https://doi.org/10.3354/cr022013.

Hitchens, N. M., H. E. Brooks, and R. S. Schumacher, 2013: Spatial and temporal characteristics of heavy hourly rainfall in the United States. *Mon. Wea. Rev.*, **141**, 4564–4575, https://doi.org/10.1175/MWR-D-12-00297.1.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.

Hope, P., and Coauthors, 2014: A comparison of automated methods of front recognition for climate studies: A case study in Southwest Western Australia. *Mon. Wea. Rev.*, **142**, 343–363, https://doi.org/10.1175/MWR-D-12-00252.1.

Johns, R. H., 1984: A synoptic climatology of northwest-flow severe weather outbreaks. Part II: Meteorological parameters and synoptic patterns. *Mon. Wea. Rev.*, **112**, 449–464, https://doi.org/10.1175/1520-0493(1984)112<0449:ASCONF>2.0.CO;2.

Jolliffe, I. T., and D. B. Stephenson, Eds., 2011: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. 2nd ed. John Wiley and Sons, 292 pp.

Kain, J. S., and Coauthors, 2008: Some practical considerations for the first generation of operational convection-allowing NWP: How much resolution is enough? *Wea. Forecasting*, **23**, 931–952, https://doi.org/10.1175/WAF2007106.1.

——, and Coauthors, 2010a: Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting–research environment. *Wea. Forecasting*, **25**, 1510–1521, https://doi.org/10.1175/2010WAF2222405.1.

——, S. R. Dembek, S. J. Weiss, J. L. Case, J. J. Levit, and R. A. Sobash, 2010b: Extracting unique information from high-resolution forecast models: Monitoring selected fields and phenomena every time step. *Wea. Forecasting*, **25**, 1536–1542, https://doi.org/10.1175/2010WAF2222430.1.

Katurji, M., B. Noonan, P. Zawar-Reza, T. Schulmann, and A. Sturman, 2015: Characteristics of the springtime Alpine valley atmospheric boundary layer using self-organizing maps. *J. Appl. Meteor. Climatol.*, **54**, 2077–2085, https://doi.org/10.1175/JAMC-D-14-0317.1.

Kennedy, A. D., X. Dong, and B. Xi, 2016: Cloud fraction at the ARM SGP site: Reducing uncertainty with self-organizing maps. *Theor. Appl. Climatol.*, **124**, 43–54, https://doi.org/10.1007/s00704-015-1384-3.

Kincer, J. B., 1916: Daytime and nighttime precipitation and their economic significance. *Mon. Wea. Rev.*, **44**, 628–633, https://doi.org/10.1175/1520-0493(1916)44<628:DANPAT>2.0.CO;2.

King, A. T., and A. D. Kennedy, 2019: North American supercell environments in atmospheric reanalyses and RUC-2. *J. Appl. Meteor. Climatol.*, **58**, 71–92, https://doi.org/10.1175/JAMC-D-18-0015.1.

Kohonen, T., 1989: *Self-Organization and Associative Memory*. 3rd ed. Springer-Verlag, 312 pp.

——, J. Hynninen, J. Kangas, and J. Laaksonen, 1996: SOM_PAK: The self-organizing map program package. Espoo: Helsinki University of Technology, Laboratory of Computer and Information Science, Rep. A31, 27 pp.

Kolczynski, W. C., and J. P. Hacker, 2014: The potential for self-organizing maps to identify model error structures. *Mon. Wea. Rev.*, **142**, 1688–1696, https://doi.org/10.1175/MWR-D-13-00189.1.

Kruschke, J. K., 2013: Bayesian estimation supersedes the t test. *J. Exp. Psychol. Gen.*, **142**, 573–603, https://doi.org/10.1037/a0029146.

Lamb, H. H., 1959: Our changing climate, past and present. *Weather*, **14**, 299–318, https://doi.org/10.1002/j.1477-8696.1959.tb00533.x.

Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Wea. Rev.*, **136**, 3408–3424, https://doi.org/10.1175/2008MWR2332.1.

Legates, D. R., and G. J. McCabe, 1999: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.*, **35**, 233–241, https://doi.org/10.1029/1998WR900018.

Lin, Y. 2011: GCIP/EOP surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data, version 1.0. UCAR/NCAR–Earth Observing Laboratory, accessed 15 November 2018, https://doi.org/10.5065/D6PG1QDD.

——, and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1.2, https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.

Liu, Y., and R. H. Weisberg, 2011: A review of self-organizing map applications in meteorology and oceanography. *Self-Organizing Maps: Applications and Novel Algorithm Design*, J. I. Mwasiagi, Ed., InTech, 253–272.

——, ——, and C. N. K. Mooers, 2006: Performance evaluation of the self-organizing map for feature extraction. *J. Geophys. Res.*, **111**, C05018, https://doi.org/10.1029/2006JA011890.

Maddox, R. A., L. R. Hoxit, C. F. Chappell, and F. Caracena, 1978: Comparison of meteorological aspects of the Big Thompson and Rapid City flash floods. *Mon. Wea. Rev.*, **106**, 375–389, https://doi.org/10.1175/1520-0493(1978)106<0375:COMAOT>2.0.CO;2.

——, C. F. Chappell, and L. R. Hoxit, 1979: Synoptic and meso-α scale aspects of flash flood events. *Bull. Amer. Meteor. Soc.*, **60**, 115–123, https://doi.org/10.1175/1520-0477-60.2.115.

Mahoney, K. M., and G. M. Lackmann, 2007: The effect of upstream convection on downstream precipitation. *Wea. Forecasting*, **22**, 255–277, https://doi.org/10.1175/WAF986.1.

Marchand, R., N. Beagley, S. E. Thompson, T. P. Ackerman, and D. M. Schultz, 2006: A bootstrap technique for testing the relationship between local-scale radar observations of cloud occurrence and large-scale atmospheric fields. *J. Atmos. Sci.*, **63**, 2813–2830, https://doi.org/10.1175/JAS3772.1.

——, ——, and T. P. Ackerman, 2009: Evaluation of hydrometeor occurrence profiles in the multiscale modeling framework

climate model using atmospheric classification. *J. Climate*, **22**, 4557–4573, https://doi.org/10.1175/2009JCLI2638.1.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430, https://doi.org/10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.*, **20**, 851–875, https://doi.org/10.1029/RG020i004p00851.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, https://doi.org/10.1175/BAMS-87-3-343.

Mitchell, M. J., R. W. Arritt, and K. Labas, 1995: A climatology of the warm season Great Plains low-level jet using wind profiler observations. *Wea. Forecasting*, **10**, 576–591, https://doi.org/10.1175/1520-0434(1995)010<0576:ACOTWS>2.0.CO;2.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the long wave. *J. Geophys. Res.*, **102**, 16 663–16 682, https://doi.org/10.1029/97JD00237.

Nelson, B. R., O. P. Prat, D. Seo, and E. Habib, 2016: Assessment and implications of NCEP Stage IV quantitative precipitation estimates for product intercomparisons. *Wea. Forecasting*, **31**, 371–394, https://doi.org/10.1175/WAF-D-14-00112.1.

Nesbitt, S. W., and E. J. Zipser, 2003: The diurnal cycle of rainfall and convective intensity according to three years of TRMM measurements. *J. Climate*, **16**, 1456–1475, https://doi.org/10.1175/1520-0442-16.10.1456.

Nielsen, E. R., G. R. Herman, R. C. Tournay, J. M. Peters, and R. S. Schumacher, 2015: Double impact: When both tornadoes and flash floods threaten the same place at the same time. *Wea. Forecasting*, **30**, 1673–1693, https://doi.org/10.1175/WAF-D-15-0084.1.

Nigro, M. A., J. J. Cassano, J. Wille, D. H. Bromwich, and M. A. Lazzara, 2017: A self-organizing-map-based evaluation of the Antarctic Mesoscale Prediction System using observations from a 30-m instrumented tower on the Ross Ice Shelf, Antarctica. *Wea. Forecasting*, **32**, 223–242, https://doi.org/10.1175/WAF-D-16-0084.1.

Nowotarski, C. J., and A. A. Jensen, 2013: Classifying proximity soundings with self-organizing maps toward improving supercell and tornado forecasting. *Wea. Forecasting*, **28**, 783–801, https://doi.org/10.1175/WAF-D-12-00125.1.

——, and E. A. Jones, 2018: Multivariate self-organizing map approach to classifying supercell tornado environments using near-storm, low-level wind, and thermodynamic profiles. *Wea. Forecasting*, **33**, 661–670, https://doi.org/10.1175/WAF-D-17-0189.1.

Nzeukou, A., H. Sauvageot, A. D. Ochou, and C. M. F. Kebe, 2004: Raindrop size distribution and radar parameters at Cape Verde. *J. Appl. Meteor.*, **43**, 90–105, https://doi.org/10.1175/1520-0450(2004)043<0090:RSDARP>2.0.CO;2.

Reusch, D. B., R. B. Alley, and B. C. Hewitson, 2007: North Atlantic climate variability from a self-organizing map perspective. *J. Geophys. Res.*, **112**, DO2104, https://doi.org/10.1029/2006JD007460.

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, https://doi.org/10.1175/2007MWR2123.1.

Root, B., P. Knight, G. Young, S. Greybush, R. Grumm, R. Holmes, and J. Ross, 2007: A fingerprinting technique for major weather events. *J. Appl. Meteor. Climatol.*, **46**, 1053–1066, https://doi.org/10.1175/JAM2509.1.

Rosenfeld, D., D. Atlas, and D. A. Short, 1990: The estimation of convective rainfall by area integrals: 2. The Height-Area Rainfall Threshold (HART) method. *J. Geophys. Res.*, **95**, 2161–2176, https://doi.org/10.1029/JD095iD03p02161.

Schumacher, R., 2017: Heavy rainfall and flash flooding. *Oxford Research Encyclopedia of Natural Hazards Science*, Oxford University Press, https://doi.org/10.1093/acrefore/9780199389407.013.132.

Schwartz, B. E., and S. Benjamin, 2000: Verification of RUC2 precipitation forecasts using the NCEP multisensor analysis. Preprints, *Fourth Symp. on Integrated Observing Systems*, Long Beach, CA, Amer. Meteor. Soc., 182–185.

Sheridan, S. C., and C. C. Lee, 2011: The self-organizing map in synoptic climatological research. *Prog. Phys. Geogr.*, **35**, 109–119, https://doi.org/10.1177/0309133310397582.

Stenz, R., X. Dong, B. Xi, and R. J. Kuligowski, 2014: Assessment of SCaMPR and NEXRAD Q2 precipitation estimates using Oklahoma Mesonet observations. *J. Hydrometeor.*, **15**, 2484–2500, https://doi.org/10.1175/JHM-D-13-0199.1.

——, ——, ——, Z. Feng, and R. J. Kuligowski, 2016: Improving satellite quantitative precipitation estimation using GOES-retrieved cloud optical depth. *J. Hydrometeor.*, **17**, 557–570, https://doi.org/10.1175/JHM-D-15-0057.1.

Stevenson, S. N., and R. S. Schumacher, 2014: A 10-year survey of extreme rainfall events in the central and eastern United States using gridded multisensor precipitation analyses. *Mon. Wea. Rev.*, **142**, 3147–3162, https://doi.org/10.1175/MWR-D-13-00345.1.

Surcel, M., I. Zawadzki, and M. K. Yau, 2016: The case-to-case variability of the predictability of precipitation by a storm-scale ensemble forecasting system. *Mon. Wea. Rev.*, **144**, 193–212, https://doi.org/10.1175/MWR-D-15-0232.1.

Tao, W.-K., and Coauthors, 2013: Precipitation intensity and variation during MC3E: A numerical modeling study. *J. Geophys. Res. Atmos.*, **118**, 7199–7218, https://doi.org/10.1002/jgrd.50410.

Tokay, A., and D. A. Short, 1996: Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds. *J. Appl. Meteor.*, **35**, 355–371, https://doi.org/10.1175/1520-0450(1996)035<0355:EFTRSO>2.0.CO;2.

Tselioudis, G., and C. Jakob, 2002: Evaluation of midlatitude cloud properties in a weather and a climate model: Dependence on dynamic regime and spatial resolution. *J. Geophys. Res.*, **107**, 4781, https://doi.org/10.1029/2002JD002259.

Van Weverberg, K., and Coauthors, 2018: CAUSES: Attribution of surface radiation biases in NWP and climate models near the U.S. Southern Great Plains. *J. Geophys. Res. Atmos.*, **123**, 3612–3644, https://doi.org/10.1002/2017JD027188.

Wallace, J. M., 1975: Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. *Mon. Wea. Rev.*, **103**, 406–419, https://doi.org/10.1175/1520-0493(1975)103<0406:DVIPAT>2.0.CO;2.

Wang, J., X. Dong, and B. Xi, 2015: Investigation of ice cloud microphysical properties of DCSs using aircraft in situ measurements during MC3E over the ARM SGP site. *J. Geophys. Res. Atmos.*, **120**, 3533–3552, https://doi.org/10.1002/2014JD022795.

——, ——, ——, and A. J. Heymsfield, 2016: Investigation of liquid cloud microphysical properties of deep convective systems: 1. Parameterization of raindrop size distribution and its application for stratiform rain estimation. *J. Geophys. Res. Atmos.*, **121**, 10 739–10 760, https://doi.org/10.1002/2016JD024941.

——, ——, and ——, 2018: Investigation of liquid cloud microphysical properties of deep convective systems: 2. Parameterization of raindrop size distribution and its application for convective rain estimation. *J. Geophys. Res. Atmos.*, **123**, 11 637–11 651, https://doi.org/10.1029/2018JD028727.

Wang, S., and T. Chen, 2009: The late-spring maximum of rainfall over the U.S. Central Plains and the role of the low-level jet. *J. Climate*, **22**, 4696–4709, https://doi.org/10.1175/2009JCLI2719.1.

Wash, C. H., S. H. Heikkinen, C. Liou, and W. A. Nuss, 1990: A rapid cyclogenesis event during GALE IOP 9. *Mon. Wea. Rev.*, **118**, 234–257, https://doi.org/10.1175/1520-0493(1990)118<0234:ARCEDG>2.0.CO;2.

Weaver, S. J., and S. Nigam, 2008: Variability of the Great Plains low-level jet: Large-scale circulation context and hydroclimate impacts. *J. Climate*, **21**, 1532–1551, https://doi.org/10.1175/2007JCLI1586.1.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.*, **125**, 527–548, https://doi.org/10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2.

——, C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23**, 407–437, https://doi.org/10.1175/2007WAF2007005.1.

Weng, S.-P., 2000: A new perspective on the regional hydrologic cycle over North and South America. Ph.D. dissertation, Iowa State University, 153 pp.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.

Wu, D., X. Dong, B. Xi, Z. Feng, A. Kennedy, G. Mullendore, M. Gilmore, and W.-K. Tao, 2013: Impacts of microphysical scheme on convective and stratiform characteristics in two high precipitation squall line events. *J. Geophys. Res. Atmos.*, **118**, 11 119–11 135, https://doi.org/10.1002/jgrd.50798.

Yang, S., and E. A. Smith, 2006: Mechanisms for diurnal variability of global tropical rainfall observed from TRMM. *J. Climate*, **19**, 5190–5226, https://doi.org/10.1175/JCLI3883.1.

Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338, https://doi.org/10.1175/2011BAMS-D-11-00047.1.